博弈论



与无线传感器网络安全

GAME THEORY
MEETS WIRELESS SENSOR
NETWORKS SECURITY

沈士根 刘建华 曹奇英 著

清华大学出版社

博弈论与无线传感器网络安全

Game Theory Meets Wireless Sensor Networks Security

沈士根　刘建华　曹奇英　著

清华大学出版社
北　京

内 容 简 介

本书以博弈论为理论分析工具，主要论述和分析无线传感器网络安全领域的若干关键问题。第 1 章介绍研究背景；第 2 章概述相关的博弈类型；第 3 章给出基于信号博弈的无线传感器网络入侵检测模型，确定何时启动入侵检测系统的最优策略；第 4 章描述基于演化博弈的无线传感器网络节点的信任模型，阐明节点信任演化动力学规律；第 5 章基于微分博弈给出无线传感器网络恶意程序传播的最优控制策略；第 6 章基于随机博弈和 Markov 链建立受攻击无线传感器网络可生存性模型，形成可生存性分析的理论和方法；第 7 章针对受攻击协调器节点，给出基于随机博弈的防御技术，再利用演化博弈实现协调器节点的选择；第 8 章阐述传感云数据外包中心访问控制系统的安全分析框架，给出基于证书认证博弈的安全优化策略；第 9 章基于随机演化联盟博弈给出受攻击虚拟传感云服务系统的自适应防御策略；第 10 章介绍无线传感器网络物理层安全技术，基于演化博弈中的复制动力学方程实现一种传感器节点保密率自适应调节的方法。

本书可作为高等院校、科研院所等从事网络安全、博弈论应用等研究人员的参考书。

图书在版编目(CIP)数据

博弈论与无线传感器网络安全/沈士根，刘建华，曹奇英著．—北京：清华大学出版社，2016
ISBN 978-7-302-41906-8

Ⅰ．①博…　Ⅱ．①沈…　②刘…　③曹…　Ⅲ．①博弈论－应用－无线电通信－传感器－安全管理－研究　Ⅳ．①TP212

中国版本图书馆 CIP 数据核字(2015)第 259856 号

责任编辑：闫红梅　赵晓宁
封面设计：
责任校对：梁　毅
责任印制：

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn，http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　**邮　　编**：100084
社 总 机：010-62770175　　**邮　　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
课件下载：http://www.tup.com.cn，010-62795954
印 刷 者：
装 订 者：
经　　销：全国新华书店
开　　本：185mm×260mm　　**印　　张**：13　　**字　　数**：323 千字
版　　次：2016 年 3 月第 1 版　　**印　　次**：2016 年 3 月第1 次印刷
印　　数：1～　000
定　　价：　.00 元

产品编号：066929-01

前 言

FOREWORD

无线传感器网络(Wireless Sensor Networks)由大量部署在监测区域内的廉价微型传感器节点组成,通过无线网络通信传输方式形成一个自组织、自适应、多跳的智能网络系统,其目的是协作地实时监测、感知和采集各种环境或监测对象的信息(如温度、湿度、气压等),再通过基站发送给管理者。当前,其在工农业、城市管理、生物医疗、环境监测、军事等众多领域已被公认具有十分广阔的应用前景。

无线传感器网络作为一种任务型网络,不仅要进行数据的传输,而且要进行数据融合、任务的协同控制等。如何保证任务执行的机密性、数据产生的可靠性以及数据传输的安全性,就成为无线传感器网络安全需要全面考虑的内容。可以说,安全问题是制约无线传感器网络发展和应用的一个关键因素。

博弈论是研究两个或多个参与者谋略和决策问题的理论,能为无线传感器网络安全的研究提供新颖的思路。无线传感器网络具有自组织、无控制中心、动态拓扑、资源有限等主要特点,这些特点决定了每个节点在通信时都会有自己的决策。那么,当节点需要做出决策时,哪一种是最优的?节点也许会表现自私而寻求只对自身有益的决策,甚至会表现恶意而选择破坏网络性能的决策。利用非合作博弈对这些情况进行研究能找到很好的答案。当然,这里的非合作博弈包括了多种形式,如信号博弈、随机博弈、微分博弈等。另外,还可以选择演化博弈对节点行为的动态演化进行研究。

本书以博弈论为理论分析工具,主要分析和解决无线传感器网络安全领域的若干关键问题。

第 1 章从无线传感器网络安全的需求出发,说明博弈论与无线传感器网络安全之间的相互关系。

第 2 章给出博弈论的基本概念,介绍适合不同情况的博弈类型,为后续章节博弈论的应用和相关工作的比较提供知识准备。

第 3 章应用信号博弈描述并分析恶意传感器节点和无线传感器网络入侵检测系统之间的交互过程。在每个独立的阶段,建立“阶段入侵检测博弈”模型,分别得到该模型的纯策略贝叶斯均衡和混合策略贝叶斯均衡。随着博弈的重复,通过构建“多阶段动态入侵检测博弈”来反映恶意传感器节点和入侵检测代理之间的交互活动,得到相应的完美贝叶斯均衡,再在此基础上实现入侵检测启动最优策略的机制和算法。

第 4 章利用演化博弈研究和分析传感器节点间的信任决策过程,根据各个传感器节点能选择不同策略的实际情况建立“无线传感器网络信任博弈”模型,通过整合激励机制参数来说明激励机制对传感器节点选择动作的影响,使用复制动态动力学方程探索博弈模型的

演化稳定策略，从而揭示无线传感器网络中各传感器节点间的信任演化原理。

第 5 章扩展经典流行病理论使之适合无线传感器网络恶意程序传播现状，并引入不同的参数来揭示无线传感器网络恶意程序传播过程。然后将恶意程序在无线传感器网络传播时“无线传感器网络系统”和“恶意程序”之间的决策交互过程看作优化控制问题，建立相应的微分博弈模型，在“恶意程序”动态改变其策略的前提下，得到“无线传感器网络系统”的最优控制策略，为控制无线传感器网络恶意程序传播的机制设计提供理论基础。

第 6 章从可靠度和可用度两方面评估受攻击无线传感器网络的可生存性属性。由于恶意攻击者总是故意发动恶意攻击行为，通过随机博弈给出这些理性恶意攻击者采取恶意攻击的期望概率，将聚簇无线传感器网络看作一个串—并系统，再利用连续时间马尔可夫链对受攻击传感器节点生命期的所有状态建立模型，基于可靠性理论得到计算受攻击传感器节点平均无故障时间、可靠度、生存期和稳态可用度的计算公式，实现受攻击无线传感器网络的可生存性评估。

第 7 章以最小化从源到目的节点的数据包分发平均跳数并且延长网络生命期为目标，提出了基于博弈论和模糊逻辑的协调器节点选择算法。在此算法中，先使用随机博弈对攻击进行动态响应，然后通过模糊逻辑选择通信质量较好的节点作为协调器节点，提高网络的服务质量和安全性。

第 8 章阐述了基于动态证书博弈的认证系统框架。在证书认证博弈交互过程中，通过认证代理补偿一定的信任度来激励传感云用户出示更多的证书，以提高其信任度。传感云用户和认证协调器通过平衡证书泄露和信任补偿之间的关系来决定是否用户能够操作外包数据。其中，认证协调器决定每次博弈信任度，认证代理决定信任度分配，再将动态证书博弈系统框架模型化为三阶段博弈，并使用迭代博弈学习方法证明信任协同的稳定性。与传统的基于属性和本体的访问控制系统相比，基于动态证书博弈的认证系统框架提高了安全效用和认证性能。

第 9 章提出了基于随机演化联盟博弈框架的受攻击虚拟传感云服务系统安全机制。在博弈的每一阶段，虚拟传感云服务提供者能够观察到服务组合节点的虚拟容量和攻击者采取的攻击策略，根据这些观察，决定需分配的虚拟容量值来保证可靠安全的服务组合。虚拟传感云服务提供者通过 minimax-Q 和演化联盟形成算法，自适应地变化防御策略，形成可靠安全的服务组合对攻击者进行动态防御。与随机博弈和演化联盟博弈相比，随机演化联盟博弈框架在动态虚拟的安全服务组合过程中获得了较好的性能。

第 10 章通过扩展经典窃听信道模型，针对聚簇无线传感器网络提出了传感器节点和其对应簇头节点之间的保密率计算方法，构建了一个非合作保密率博弈模型，以反映传感器节点之间的交互关系。利用演化博弈思想，建立了传感器节点自适应选择发射功率的机制，提出了传感器节点保密率的自适应调节算法，为保证无线传感器网络数据的保密性提供了新途径。

本书是作者多年研究博弈论和无线传感器网络安全的成果，其中，沈士根教授负责撰写第 1～6 章和第 10 章，刘建华博士负责撰写第 7～9 章，曹奇英教授负责统稿及理论指导。

作者的研究工作得到了国家自然科学基金项目(61272034,61572014)、浙江省自然科学基金项目(LY16F020028)、中央财政《无线 Mesh 网络若干关键技术研究创新团队建设项目》等项目的资助。在本书的撰写过程中,绍兴文理学院机械与电气工程学院、嘉兴学院数理与信息工程学院、东华大学计算机科学与技术学院给予了大力支持,在此一并表示感谢。

由于作者水平所限,加之博弈论在网络安全中的应用研究处于不断发展和变化之中,书中错误和不足之处在所难免,恳请专家、读者予以指正。

作　者

2015 年 8 月

目录

CONTENTS

第 1 章

绪　论

本章从无线传感器网络的研究背景和无线传感器网络安全的需求出发，说明博弈论与无线传感器网络安全之间的关系，给出本书的组织结构。

1.1 研究背景

微电子技术、计算技术和无线网络通信等技术的发展，促进了低功耗多种类传感器的快速发展，使其在微小体积内能够实现信息收集、数据计算和无线网络传输等多种功能。无线传感器网络（Wireless Sensor Networks）就是由大量部署在监测区域内的廉价微型传感器节点组成的，通过无线网络传输方式形成的一个多跳的自组织、自适应的智能网络系统，其功能是合作地感知、收集并处理网络覆盖区域中各类对象（如温度、湿度、气压等）的信息，再发送给管理者。因此，组成一个传感器网络的 3 个主要要素是传感器节点、感知对象和管理者。如果说因特网构成了逻辑上的信息世界，改变了人与人之间的沟通方式，那么，无线传感器网络就是将客观上的物理世界与逻辑上的信息世界融合在一起，改变人类与自然界的交互方式。人们可以通过无线传感器网络直接感知物理世界中各类对象信息，从而极大地扩展现有网络的功能和人类认识物理世界的能力。美国商业周刊和 MIT 技术评论在预测未来技术发展的报告中，分别将无线传感器网络列为 21 世纪最有影响力的 21 项技术和改变世界的十大技术之一[1]。研究结果[2-5]表明，无线传感器网络具有十分广阔的应用前景，在工农业、城市管理、生物医疗、环境监测、军事等众多领域都有实际与潜在的实用价值。

无线传感器网络经历了一个长期的发展过程。在 20 世纪 70 年代，出现的第一代传感器网络主要利用点对点传输技术以及专门的控制器将传统的传感器连接起来，从而形成了无线传感器网络的雏形。随后，电子、计算机等学科的不断发展和进步，使传感器网络也具备了获取多种对象信息的综合处理能力，并采用串/并接口与传感控制器相连，构成了具有信息收集和综合处理能力的第二代传感器网络。第三代传感器网络形成于 20 世纪 90 年代后期和 21 世纪初，开始采用能够智能获取多种对象信息信号的传感器，通过现场总线连接传感控制器，形成局部智能化传感器网络。第四代传感器网络是目前科研工作者的研究热点之一，该网络采用大量具有多功能、多对象信号获取能力的传感器，尤其重要的变化是传感器之间采用可靠的无线网络传输协议进行连接，从而形成高效、健壮的无线传感器网络，这是传感器网络发展的一个巨大飞跃[6]。这将使传感器网络进一步发展，应用范围得到极大的扩展。

从科研的角度来看,无线传感器网络的研究起始于20世纪90年代末期。自1999年将中间件(Middleware)技术引入无线传感器网络中之后,就有很多科研院所开始从不同的侧面进行研究。那时,大多数开展的基于无线传感器网络特性的中间件研究和开发工作都主要集中在如何延长传感器网络的生命期以及如何充分提高传感器网络的有限资源利用等方面。在美国,康奈尔大学、加州大学伯克利分校等是较早开始无线传感器网络基础理论和关键技术研究的高校。此后,大家都认识到无线传感器网络具有巨大的实际应用价值,世界许多国家的军事部门、工业界和学术界都对这种网络表现出极大的关注。美国自然科学基金委员会(US National Science Foundation)于2003年制订了无线传感器网络的研究计划,大力支持无线传感器网络基础理论和关键技术的研究。由于无线传感器网络潜在的军事用途,美国国防部(US Department of Defense)对此也高度重视,把无线传感器网络作为一个重要的研究领域,设立了一系列的项目从事军事传感器网络的研究;美国英特尔(Intel)公司、微软(Microsoft)公司等信息业巨头也开始了无线传感器网络方面的研究工作;其他如意大利、俄罗斯、法国、日本、英国、德国等科技发达国家也对无线传感器网络表现出了极大的兴趣,纷纷展开了相关的科学研究工作[7]。

我国的中国科学院上海微系统研究所、计算所、软件研究所、沈阳自动化所、电子所和合肥智能技术研究所等科研机构,清华大学、北京大学、哈尔滨工业大学、西北工业大学、北京邮电大学、南京邮电大学、国防科技大学等高等院校在国内较早开展了传感器网络的研究,之后有更多的科研院所加入到无线传感器网络的基础研究和开发工作中来。

通常,典型的无线传感器网络包括传感器节点(Sensor Node)、汇聚节点(Sink Node)和管理节点[5]。大量的传感器节点以随机撒播的方式部署在监测区域内部或附近,能够通过自组织的方式互联成网络。各类传感器节点监测到的数据信息沿着其他传感器节点(如簇头)逐跳地进行传输,并在传输过程中不同节点的监测数据信息可能被多个节点进行处理,再经过多跳后传递到汇聚节点,最后通过互联网传输到管理节点。管理者可通过管理节点对传感器网络进行管理和配置,收集监测数据和发布监测信息等任务[8,9]。

但由于无线传感器网络感知、收集和传输数据的性能受到环境和节点自身特点的约束,在实际应用中存在诸多不足之处,主要体现在以下几个方面。

1. 电源能量有限

传感器节点体积微小,通常携带能量十分有限的电池[10]。这些能量主要被传感器模块、处理器模块和无线通信模块等消耗。随着集成电路工艺的发展,传感器和处理器模块的功率消耗将会变得越来越低,绝大部分能量消耗在无线通信模块上。其中,无线通信模块具有接收、发送、睡眠、空闲4种状态。空闲状态意味着无线通信模块一直在监听无线信道的状况,检查是否有数据信息发送过来,而睡眠状态则意味着关闭无线通信模块。相比较而言,无线通信模块在数据发送时能量消耗最大,空闲时少于发送状态的能量消耗,而处于睡眠状态时能量消耗最少[1]。由于一个无线传感器网络中的传感器节点个数多、分布区域广,而且部署环境复杂,有些部署区域甚至人员都不能到达,所以通过更换电池的方式来补充能源往往不现实。这就对科研工作者提出了无线传感器网络多方面节能的需求。

2. 通信能力有限

传感器节点能量有限的现状决定了它有限的通信能力。无线网络通信的能耗与通信距离的关系密切,随着通信距离的增加,能量消耗将成倍增加。考虑到传感器节点网络覆盖区

域大的特点,无线传感器网络通常采用多跳路由传输机制。这就要求在满足无线传感器网络通信连通度的前提下应尽量减少单跳通信距离。另外,由于节点能量的不断变化,受障碍物等自然环境的影响,无线网络通信性能会经常变化,导致通信中断[1]。这就对科研工作者提出了多方面减少数据通信的需求。

3. 计算和存储能力有限

作为一种微型嵌入式设备,传感器节点价格低、功耗小,这些限制必然导致其配备的微处理器能力比较弱,存储器容量比较小。而传感器节点需要完成监测数据的采集和转换、数据的管理和处理、应答汇聚节点的任务请求和节点控制等多种工作[1]。为了完成各种任务,这就对科研工作者提出了多方面减少数据计算和存储的需求。

因此,目前针对无线传感器网络的大量科研工作都是为了解决上述的不足进行展开。

实际上,无线传感器网络作为任务型的网络,不仅要进行数据的传输,而且要进行数据采集和融合、任务的协同控制等。如何保证任务执行的机密性、数据产生的可靠性、数据融合的高效性以及数据传输的安全性,就成为无线传感器网络安全需要全面考虑的内容。可以说,安全问题是制约无线传感器网络发展的一个非常关键因素[11-14]。

博弈论是研究两个或多个参与者谋略和决策问题的理论[15],在我国古代故事如王戎辩李、孙膑赛马、破釜沉舟、空城计等中就充满了博弈论的思想。博弈论分析的目的是预测博弈的结果。不言而喻,每一个参与者要选择的策略必须是针对其他参与者选择战略的最优反应,每一个参与者都希望尽可能提高自己的利益所得[16]。因此,博弈论研究强调决策主体行为发生时的直接相互作用。例如,经常遇到的性别战博弈,这是一个两人决策问题,丈夫的决策依赖于妻子的决策;反过来,妻子的决策也依赖于丈夫的决策。

1944 年,John Von Neumann 和 Oskar Morgenstern 的巨著 *Theory of Games and Economic Behavior* 的出版为博弈论在经济学中系统的应用奠定了理论基础。1994 年、1996 年、2001 年、2005 年和 2007 年共 5 次诺贝尔经济学奖被分别授予了经济博弈论方向的学者。对一门学科给予如此高的褒奖,表明了博弈论的强大威力。正是国际经济学界对博弈论的这种肯定,推动了国内外博弈论研究及应用的发展,使得目前的博弈论已发展成一个内容丰富且完善的理论体系。更重要的是,博弈论的应用已逐步扩展到政治学、道德哲学、社会学、生物学和计算机科学等领域。

博弈论的应用需要根据不同的场合选择不同的博弈类型,本书主要利用非合作博弈、演化博弈、联盟博弈研究无线传感器网络安全中的若干关键问题。非合作博弈的核心问题是参与者的策略选择,即在参与人是完全理性的基础上研究参与者在利益相互影响的情况下选择最有利于自己的策略[17]。演化博弈建立在参与人是有限理性的基础上,以参与人种群为研究对象,认为参与人的行为是一个动态调整过程[18]。联盟博弈强调在联盟的内部建立信息的互通,以及具有约束力且可执行的契约。因此,非合作博弈适用于参与人存在竞争且需要探寻只对自身有利策略的场合,演化博弈适用于需要对参与人行为动态演化进行研究的场合,而联盟博弈适用于联盟是否可获得收益,以及获得的净收益如何在联盟内部公平分配的问题。

近些年来,在无线网络领域,包括 Ad Hoc 网络、Mesh 网络、无线传感器网络等,博弈论的应用呈明显上升的趋势。研究涉及的内容包括无线传感器网络媒体接入控制、无线传感器网络安全路由、无线传感器网络 MAC 协议竞争接入控制、认知 MIMO 系统功率分配、毫

微微蜂窝混合接入控制干扰管理、高速移动环境下快速动态无线资源优化、无线自组织网络用户合作激励、认知无线网络动态频谱拍卖、认知无线网络资源分配、智能绿色无线电资源分配、认知无线电动态频谱分配等。

国内外一些著名研究机构和学术团队都在致力于博弈论和无线网络相结合的研究，如中国科学院软件研究所信息安全国家重点实验室、上海交通大学系统控制与信息处理教育部重点实验室、东北大学计算机软件与理论研究所、华东交通大学智能传感器网络中心和网络与信息安全中心、武汉理工大学高性能网络研究中心、四川大学计算机网络与安全研究所、西安电子科技大学智能感知与图像理解实验室、哈尔滨工业大学通信技术研究所、美国佐治亚理工学院宽带无线网络实验室、美国马里兰大学 K. J. Ray Liu 团队、美国加州大学 Mihaela van der Schaar 团队、美国伊利诺大学 Tamer Başar 团队、美国范德堡大学 Myrna Wooders 团队、加拿大曼尼托巴大学 Ekram Hossain 团队、希腊西马其顿大学 Athanasios V. Vasilakos 团队等。与此同时，从近几年的网络通信领域顶级国际会议 SIGCOMM、INFOCOM 和 MobiHoc 上发表的论文可见，每年都有相关文献发表。通信领域国际顶级期刊 *IEEE Journal on Selected Areas in Communications* 在 2011 年的征稿中共有两次主要关注博弈论和无线网络相结合的主题，分别是 Game Theory in Wireless Communications 和 Economics of Communication Networks and Systems。

事实上，博弈论为无线传感器网络安全的研究提供了新颖的思路。自组织、无控制中心、动态拓扑、资源有限是无线传感器网络的主要特点，这些特点决定了每一个节点在通信时会有自己的决策。那么，当节点需要做出决策时，哪一种是最优的？节点也许会表现自私而寻求只对自身有益的决策，甚至会表现恶意而选择破坏网络性能的决策。这些情况利用非合作博弈能找到很好的答案。当然，这里的非合作博弈包括了多种形式，如重复博弈、信号博弈、声明博弈、随机博弈等。另外，还可以选择演化博弈对节点行为的动态演化进行研究。因此，博弈论方法为无线传感器网络安全中多方面关键问题研究提供了可行的新思路和新技术，这是一个重要的充满前景的研究方向。

1.2 本书组织结构

本书总共包含 10 章，分别是“第 1 章 绪论”、“第 2 章 博弈论概述”、“第 3 章 基于信号博弈的无线传感器网络入侵检测最优策略研究”、“第 4 章 基于演化博弈的无线传感器网络节点信任演化动力学研究”、“第 5 章 基于微分博弈的无线传感器网络恶意程序传播机制研究”、“第 6 章 基于随机博弈的受攻击无线传感器网络可生存性评估研究”、“第 7 章 无线传感器网络受攻击协调器节点的防御响应博弈机制研究”、“第 8 章 面向传感云数据外包中心的信任演化机制研究”、“第 9 章 基于随机演化联盟博弈的虚拟传感云服务安全机制研究”、“第 10 章 基于演化博弈的传感器节点保密率自适应调节研究”。

第 1 章由无线传感器网络的研究背景和无线传感器网络安全的需求，说明博弈论与无线传感器网络安全之间的关系。

第 2 章概要介绍了博弈论的基本概念、博弈类型等。通过对完全信息静态博弈、完全且完美信息动态博弈、重复博弈、不完全信息静态博弈、完全但不完美信息动态博弈、不完全信息动态博弈、合作博弈、信号博弈、演化博弈、微分博弈、随机博弈、联盟博弈等博弈类型的说

明和分析,初步了解博弈论,为后续章节博弈论的应用和相关工作的比较提供知识准备。

第 3 章应用信号博弈描述并分析恶意传感器节点和无线传感器网络入侵检测系统之间的交互过程。在每个独立的阶段,建立“阶段入侵检测博弈”(Stage Intrusion Detection Game)模型,探索该博弈模型纳什均衡存在的条件,并将分别得到纯策略贝叶斯均衡(Pure-strategy Beyasian Equilibrium)和混合策略贝叶斯均衡(Mixed-strategy Beyasian Equilibrium)。随着博弈的进行,通过构建“多阶段动态入侵检测博弈”(Multi-stage Dynamic Intrusion Detection Game)来反映恶意传感器节点和入侵检测代理之间的交互活动。另外,在得到“多阶段动态入侵检测博弈”的完美贝叶斯均衡的基础上,设计入侵检测运行机制和相应的算法。

第 4 章利用演化博弈研究传感器节点间的信任决策过程,从而揭示无线传感器网络各传感器节点间的信任演化原理。根据各个传感器节点能选择不同策略的实际情况建立“无线传感器网络信任博弈”模型,并且为了研究激励机制对传感器节点选择动作“信任”(即可以合作通信)的影响,在“无线传感器网络信任博弈”模型中整合激励机制参数。为了说明“无线传感器网络信任博弈”模型的稳定性,通过复制动态动力学方程探索“无线传感器网络信任博弈”的演化稳定策略。

第 5 章扩展经典流行病理论使之适合无线传感器网络恶意程序传播现状,并引入不同的参数来揭示无线传感器网络恶意程序传播过程。然后将恶意程序在无线传感器网络传播时“无线传感器网络系统”和“恶意程序”之间的决策问题看作优化控制问题,并利用微分博弈为“无线传感器网络系统”得到最优控制策略,这种策略将在考虑“恶意程序”最大化破坏无线传感器网络的前提下,最小化“无线传感器网络系统”和“恶意程序”产生的成本。

第 6 章从可靠度和可用度两方面评估受攻击无线传感器网络的可生存性属性。首先将选择研究的聚簇无线传感器网络看作一个串-并(Serial-parallel)系统,这样就可以应用经典可靠性理论中已有的结论。因为恶意攻击者总是故意发动恶意攻击行为,通过随机博弈给出这些理性恶意攻击者采取恶意攻击的期望概率,再利用连续时间马尔可夫链对受攻击传感器节点生命期的所有状态建立模型,就可得到计算受攻击传感器节点平均无故障时间、可靠度、生存期(Survival Lifetime)和稳态可用度的计算公式,实现受攻击无线传感器网络的可生存性评估。

第 7 章针对受攻击的 ZigBee 无线传感器网络,为了最小化从源到目的节点数据包分发的平均跳数并且最大化延长网络生命期,提出了基于博弈论和模糊逻辑的协调器节点选择算法。协调器节点选择算法不仅要考虑数据包分发延迟、网络生命期,而且还要考虑网络攻击防御策略,因此在提出的算法中,首先使用随机博弈对攻击进行动态响应,然后通过模糊逻辑选择服务质量较好的节点作为协调器节点,从而提高网络的服务质量和安全性。

第 8 章阐述了基于证书认证的信任演化博弈模型及其系统框架。在“证书认证信任演化博弈”交互过程中,通过认证代理补偿一定的信任度来激励传感云用户披露更多的证书,以提高其信任度。传感云用户和认证协调器通过平衡证书泄露和信任补偿之间的关系来决定用户是否能够执行外包数据访问操作,认证代理根据用户披露的证书决定信任度的分配,并使用多轮迭代博弈效用分析法分析了“证书认证信任演化博弈”的稳定性。与传统的基于属性和本体的访问控制系统相比,基于证书认证的信任演化博弈模型及其系统框架提高了安全效用和认证性能。

第 9 章提出了基于随机演化联盟博弈的受攻击虚拟传感云服务系统安全机制。在随机演化联盟博弈的每一阶段，虚拟传感云服务提供者能够观察到服务组合节点的虚拟容量和攻击者的策略，根据这些观察，决定需分配的虚拟容量值来提高虚拟传感云服务组合的服务质量。虚拟传感云服务提供者通过 minimax-Q 和演化联盟形成算法，自适应地变化其防御策略并形成可靠安全的服务组合联盟，从而对攻击者的攻击进行动态防御来提高虚拟传感云服务的安全性和可靠性。与随机博弈和演化联盟博弈相比，随机演化联盟博弈框架在虚拟传感云服务动态组合过程中获得了较好的性能。

第 10 章以最大化网络效用为目标，利用演化博弈论中的复制动力学方程实现传感器节点保密率的自适应调节机制。通过扩展经典保密率计算公式，首先建立了能适应聚簇无线传感器网络环境中簇成员传感器节点和簇头传感器节点之间的保密率计算公式。然后，通过建立一种非合作传感器节点保密率博弈模型，解决了传感器节点最大化各自保密率时影响整个网络通信的问题。最后，利用演化博弈论中的复制动力学方程，给出传感器节点如何动态地选择各自的发射功率来最大化其保密率适应度的演化过程，实现传感器节点保密率自适应调节机制。

第 2 章

博弈论概述

本章给出博弈论的基本概念，介绍适用于不同应用场合的博弈类型，包括完全信息静态博弈、完全且完美信息动态博弈、重复博弈、不完全信息静态博弈、完全但不完美信息动态博弈、不完全信息动态博弈、合作博弈、信号博弈、演化博弈、微分博弈、随机博弈、联盟博弈等，为后续章节博弈论的应用和相关工作的比较提供知识准备。

2.1 博弈论基本概念

博弈论(Game Theory)是现代数学的一个新分支，也是运筹学的重要构成内容之一。博弈论主要研究具有相互依赖行为的参与者的策略选择。现在通常所说的博弈论一般是指非合作博弈理论，认为参与者是理性的，即参与者之间都会在一定的约束条件下最大化自身的利益，同时参与者之间在交互时利益有冲突，行为相互有影响，而且不同参与者掌握的信息常常是不对称的。在这种情况下，博弈论研究参与者的行为、交互时的策略和策略的均衡问题[19, 20]。当然，现代博弈论还包括合作博弈。合作博弈强调的是团体理性、集体的效率、公正和公平[19]。

博弈论作为研究多人谋略和策略问题的理论。首先，一个博弈问题必须至少有两个参与博弈的参与者，在博弈过程中他们都有各自的切身利益。由于各自利益的驱动，他们在做出自己各自的决策时，总想使用最优策略；其次，博弈中的各个参与者之间总不可避免地存在着竞争。竞争贯穿了整个博弈的全过程，同时这种竞争又将博弈的参与者紧紧地联系在一起，相互较量，相互依存；再次，既然参与者之间要进行较量，那么每一个博弈参与者都会尽量掌握对手的特点及其已经采取或可能采取行动的相关知识和信息；最后，就是博弈参与者最为关心的博弈结果[15, 16, 21]。博弈结果随不同参与者采取策略的不同而不同，通常用支付(Payoff)来描述博弈结果。因此，博弈论就是从理论上对博弈参与者之间的行为和交互过程进行研究和分析，为博弈参与者预测出一个理想的结局。这种预测结局的正确性主要体现在博弈参与者都能自愿选择博弈理论为其推导出的策略，并且没有博弈参与者愿意独自偏离其依照博弈理论已选定的策略。所以，每个博弈参与者所选策略是针对其他参与者所选策略的最优反应。

下面介绍博弈论中的一些基本概念。

1) 参与者

参与者(Player)是指一个博弈中独立决策、独立承担后果的决策主体，通常又称为局中

人或参与人。参与者参加博弈的目的是通过各自合理地选择相应的行动,以便最大化自己的支付(或效用)水平。参与者可以是个人,也可以是具有智能管理能力的设备(如无线传感器网络中的传感器节点)、团队、企业、国家,甚至是国家组成的集团。为了研究和分析博弈问题的需要,通常还会引入一个虚拟参与者——"自然"(Nature)。这里的"自然"指的是独立于博弈参与者的外生事件。"自然"选择的是外生事件的各种可能现象,并且用概率分布来描述"自然"的选择机理[17]。

2) 信息

信息(Information)是指参与者在博弈过程中能了解到和观察到的知识,这些知识包括"自然"的选择、其他参与者的特征和行动等[17]。信息是整个博弈过程中非常重要的一个变量,一旦信息结构变化了,博弈双方的所有结果都有可能发生变化。

3) 共同知识

共同知识(Common Knowledge)是指"所有参与者知道,所有参与者知道所有参与者知道,所有参与者知道所有参与者知道所有参与者知道……"的知识[17]。

4) 完全信息

完全信息(Complete Information)是指所有参与者各自选择的策略的不同组合所决定的各参与者的收益对所有参与者来说是共同知识[17]。简单、通俗地说,完全信息是指每一个参与者完全了解自己以及其他参与者的策略,完全了解各参与者选择的策略组合产生的效用。

5) 完美信息

完美信息(Perfect Information)是指所有参与者在选择各自策略时,其他参与者的博弈进程对所有参与者而言是共同知识,也就是说,拥有完美信息就是完全了解其他参与者的行动;相反,不完美信息意味着参与者在博弈进程信息的掌握程度上具有不对称性。

6) 静态

如果在博弈过程中各参与者同时选择各自的策略,则称这类博弈是静态(Static)的。这里所说的"同时"具有双层含义:一种含义就是"同时"的字面解释,也就是不同的参与者在同一时间一起行动;另一种含义是不同的参与者行动虽然有先后,但后行动者并不知道先行动者采取了什么样的具体行动[17]。

7) 动态

动态(Dynamic)是指不同参与者的行动有先后顺序,并且后行动者能够观察到先行动者所选择的行动[17]。

8) 策略

策略(Strategy)是指参与者可选择的全部行为的集合,即规定每个参与者在进行决策时可以选择的做法。在一个静态博弈(Static Game)中,一个策略是参与者的一个给定的可能行动;在动态博弈(Dynamic Game)中,一个策略是参与者在每个决策点选择的一个完整计划,它告诉参与者在什么时候应选择什么行动[17]。

9) 支付

支付(Payoff)是指参与者在一个特定的策略组合中得到或失去的效用,它通常是参与者策略的函数,其值可以是正数也可以是负数。如果结果是随机的,那么支付通常用概率来加权平均,即期望支付(Expected Payoff)[17]。

10）均衡

均衡(Equilibrium)是指所有参与者的最优策略的组合。需要注意的是,不同的博弈类型通常具有不同的均衡形态。达到博弈均衡意味着相关量处于稳定状态,这种稳定状态在博弈过程中是可以预测的。

11）理性

如果一个参与者寻求以一种最大化自己支付的方式进行博弈,那么,这个参与者就是理性(Rationality)的[17]。以参与者个体利益最大化为目标的被称为“个体理性”,而追求集体利益最大化的被称为“集体理性”;有完美的分析判断能力和不会犯选择行为错误的称为“完全理性”,反之称为“有限理性”[19]。

12）纯策略

纯策略(Pure-strategy)是指每个参与者在博弈过程中可以选择采用的行动方案,每个参与者均有可供其选择的多种策略[17]。

13）混合策略

混合策略(Mixed-strategy)是指参与者在纯策略空间上的一种概率分布,表示参与者实际博弈时根据这种概率分布在纯策略空间中随机选择行动方案并加以实施[17]。

14）零和

如果对任何策略组合,所有参与者的支付和为零,则称该博弈是零和(Zero)的[17]。

2.2 博弈类型

2.2.1 完全信息静态博弈

在国内外学者发表的文献中,若只说明是基于博弈论方法,而未说明具体的博弈类型,那么这种博弈类型实际上就是零和的完全信息静态博弈(Complete Information Static Game)。完全信息静态博弈就是各参与者同时决策,且所有参与者对各方支付都了解的博弈。完全信息静态博弈通常使用标准式描述,包含3个方面的信息:①博弈参与者集合;②每个参与者的策略空间;③每个参与者的支付函数[16]。

与完全信息静态博弈对应的均衡就是最常说的纳什均衡[17]。纳什均衡实际上描述的是一种策略集,在这个策略集中,每一个参与者都确信,在给定对方策略的情况下,他选择了最好的策略。也就是说,参与者双方都认为自己现有的策略是最好的策略,因此,在对方不改变策略的前提下,任何一方都不会调整自己的策略;否则,率先改变策略的一方将减少对应的效用值。

判断一个结果是不是纳什均衡的通常办法是看参与者是否可以通过单方面的背离而获得更多的效用。如果还有其他的策略可以让任何一个参与者得到更多的效用,那么他一定会丢弃现在的策略组合,也就是说,现在的策略组合是不稳定的。实际上,纳什均衡是完全信息静态博弈的解,构成纳什均衡的策略一定是重复剔除严格劣策略过程中不能被剔除的策略;也就是说,没有一种策略能严格优于纳什均衡策略[19]。

需要注意的是,纳什均衡包括纯策略纳什均衡(Pure-strategy Nash Equilibrium)和混合策略纳什均衡(Mixed-strategy Nash Equilibrium)。纳什在他1950年的经典论文中,证

明了混合策略纳什均衡普遍存在于不同的博弈类型中，指出每一个有限次博弈都至少存在一个混合策略纳什均衡[15]。

2.2.2 完全且完美信息动态博弈

在完全且完美信息动态博弈(Complete and Perfect Information Dynamic Game)中，各参与者不是同时而是先后选择策略，每个参与者需要考虑如果采取这个策略，那么对方将如何应对该策略，同时还需要考虑当前采取的策略将如何影响自己及对手将来如何选择策略。这一特点使得完全且完美信息动态博弈的表示常使用扩展式博弈树描述。与支付矩阵表示法相比，扩展式博弈树扩展了参与者的策略空间，即某个参与者在什么时候行动、每次行动可选择哪些策略以及当前知道哪些信息。

完全且完美信息动态博弈中各个参与者策略选择的先后顺序形成了连续的博弈过程，其中各参与者的一次选择行为称为一个“阶段”。如果完全且完美信息动态博弈的几个参与者同时选择策略，那么这些参与者的同时选择也构成一个“阶段”。一个完全且完美信息动态博弈至少包含两个“阶段”，因此常把完全且完美信息动态博弈也称为“多阶段博弈”[19]。完全且完美信息动态博弈也被称为“序贯博弈”，这是从各参与者选择策略有时间先后方面进行考虑的。完全且完美信息动态博弈还被称为“扩展博弈”，这是因为完全且完美信息动态博弈常采用扩展式博弈树来表示各参与者的选择次序和各博弈阶段。

一个完全且完美信息动态博弈包含 6 个方面的信息：①博弈参与者集合；②参与者的行动顺序；③每次轮到某参与者行动时，可供他选择的行动；④每次轮到某参与者行动时，他了解到的信息；⑤各个参与者选择不同的行动组合后对应的支付；⑥虚拟参与者“自然”可能选择的概率分布[16]。

与完全且完美信息动态博弈相关的均衡是“子博弈完美纳什均衡”(Subgame Perfect Nash Equilibrium)和“颤抖手完美均衡”(Trembling Hand Perfect Equilibrium)[15]。经典的模型包括斯坦克博格(Stackelberg)模型[22]、讨价还价(Bargaining)模型[23]、委托人—代理人模型[15]。其中“子博弈”由一个完全且完美信息动态博弈第一阶段后的任一阶段开始的后续“阶段博弈”构成，能够自成一个博弈，包含有初始信息集和进行博弈所需要的全部信息[19]。类似地，一个“子博弈”还可以包含下一级“子博弈”。需要注意的是，完全且完美信息动态博弈本身不是它自己的一个“子博弈”，这与集合的性质不同。另外，“子博弈”不能分割初始信息集且必须包含第一个阶段后的所有“阶段博弈”。要使一个“策略对”成为“子博弈完美纳什均衡”，必须要求它首先是原完全且完美信息动态博弈的纳什均衡，其次在完全且完美信息动态博弈的所有“子博弈”中都构成纳什均衡。与纳什均衡不同的是，“子博弈完美纳什均衡”能够排除均衡策略中不可信的威胁或承诺，排除不稳定、不合理的纳什均衡，留下真正稳定的纳什均衡[19]。而“颤抖手完美均衡”是对纳什均衡的一个改进，研究每个参与者都有可能犯错误前提下的纳什均衡。它要求参与者采用的策略，不仅在其他参与者不犯错误时是最优的，而且在其他参与者偶尔犯错误时仍然是最优的[19]。因此，“颤抖手完美均衡”是一种相当稳定的纳什均衡。

2.2.3 重复博弈

重复博弈(Repeated Game)指重复进行基本博弈而构成的博弈过程。通常研究的大多

数重复博弈是静态博弈的重复，其中的每次博弈被称为“阶段博弈”，而重复博弈又是一个动态过程，属于动态博弈的范畴，因此重复博弈与静态博弈和动态博弈都有关系[24]。虽然重复博弈形式上是原基本博弈的反复，但参与者的行动和博弈结果却不一定是原基本博弈的简单重复。如果参与者的行动在每个“阶段博弈”后都可被观察到，那么参与者就可能参考其他参与者前面的博弈行为来选择自己的策略，这样就可能导致不同的均衡结果，所以，不能简单地把重复博弈看成是原基本博弈的线性累加。

重复博弈根据重复原基本博弈的次数常可分为“有限次重复博弈”和“无限次重复博弈”[19]。显然，“有限次重复博弈”表示博弈重复次数有限，且有预定的结束时间，而“无限次重复博弈”表示无限次地重复原基本博弈。另外需要注意的是，还有一种称为“随机结束重复博弈”的重复博弈，它的博弈重复次数是有限的但博弈结束的时间和具体的博弈重复次数却是不确定的。

与独立的单次静态博弈和动态博弈不同，在重复博弈中每个参与者在每个阶段都需要进行可能不同的策略选择，这是因为各参与者在前面阶段的博弈中的策略已成为共同知识，参与者可以在此基础上进行策略选择。与动态博弈类似地是，重复博弈也有“子博弈”的概念。这些“阶段子博弈”就是从某个阶段（不包括第一阶段）开始，直到最后一个阶段的所有“阶段博弈”。与原来的重复博弈相比较，“子博弈”要么是重复次数减少的重复博弈，要么仍是原来的重复博弈（对无限次重复博弈而言）。

重复博弈的效用与单次静态博弈和动态博弈不同，它不是整个重复博弈结束后的一个总的效用，而应包含博弈过程的每个“阶段博弈”中产生的效用。对于“有限次重复博弈”，一种计算重复博弈效用的方法是累加参与者在各“阶段博弈”中的效用，简称“总效用”法；另一种方法是将总效用除以重复次数，即“平均效用”法。而对“无限次重复博弈”，由于不同时间获得的利益对参与者的价值是不相同的，因此常引入“贴现系数”将后一“阶段博弈”的效用折算成当前阶段的效用。

2.2.4　不完全信息静态博弈

不完全信息静态博弈（Incomplete Information Static Game）又称静态贝叶斯博弈，这里的不完全信息并不是完全没有信息，不完全信息静态博弈的参与者至少必须有关于其他参与者支付的可能范围和分布概率的知识；否则参与者的决策就会完全失去依据[16]。实际上，在不完全信息静态博弈中，各参与者都知道自己的效用函数，但不能确切地知道其他参与者的效用函数。另外，虽然参与者不能确定其他参与者在相应策略下的效用，但知道其他参与者有哪些可能的效用结果，而具体哪种效用结果会出现则取决于参与者属于哪种“类型”。这些“类型”是参与者自己清楚但其他参与者无法知道的个人信息，即非共同知识。因此，在求解不完全信息静态博弈时，常把博弈过程中参与者对其他参与者效用的不了解转化成对这些参与者“类型”的不了解，也就是说，在分析不完全信息静态博弈时，就必须把关注各参与者的效用转向各参与者的“类型”及采取的策略组合[19]。

不完全信息静态博弈使用标准式描述，与完全信息静态博弈不同的是它包括 5 个方面的信息：①博弈参与者集合；②参与者的类型空间；③参与者在知道自己类型的条件下，对其他参与者的类型组合推断；④依赖于类型的策略空间；⑤依赖于类型的支付函数[16]。

通过海萨尼（Harsanyi）转换引入虚拟参与者“自然”并将静态博弈赋予时间顺序，可以

把不完全信息静态博弈转化为完全信息动态博弈，然后就可以利用完全信息动态博弈的处理方法实现不完全信息静态博弈的分析。因此，不完全信息静态博弈可以看作是先由“自然”选择各参与者的类型，然后再由各参与者同时进行策略选择的动态博弈，这样不完全信息静态博弈中各参与者的一个策略实际上就是针对自己各种可能的类型如何进行选择的问题。所以，不完全信息静态博弈中参与者的策略是关于类型空间和行动空间的函数，所有的这些函数构成了参与者的策略空间[19]。

由不完全信息静态博弈得到的均衡概念称为贝叶斯均衡(Bayesian Equilibrium)。在一个有限不完全信息静态博弈中，必定存在贝叶斯均衡或混合策略贝叶斯均衡。这种贝叶斯均衡概念意味着参与者的行动是同时发生的，没有时间先后顺序，因此，没有任何参与者能够有机会观察其他参与者的选择。在给定其他参与者的策略前提下，每个参与者的最优策略实际上依赖于自己的类型。每个参与者虽然不知道其他参与者真正选择了什么策略，但只要知道其他参与者的类型的概率分布，就能够正确地预测出其他参与者的策略选择与各自类型之间的关系。所以，不完全信息静态博弈中各参与者选择策略的依据就是在给定自己类型和其他参与者的类型与策略选择之间关系的前提下，使得自己的期望支付达到最大化[19]。

2.2.5 完全但不完美信息动态博弈

完全但不完美信息动态博弈(Complete but Imperfect Information Dynamic Game)研究的博弈情况具有以下特征：①各参与者在博弈结束时完全清楚每个参与者的效用；②后行动的参与者无法或部分看到自己选择策略之前的博弈过程，或者不同的参与者掌握的博弈进程信息有差异，再或者各参与者有多次策略选择，但无法观察到前面的博弈进程[19]。完全但不完美信息动态博弈的表示仍使用扩展式博弈树描述。

与完全但不完美信息动态博弈相关的均衡概念称“完美贝叶斯均衡”(Perfect Bayesian Equilibrium)。一个“完美贝叶斯均衡”必须要满足以下要求：①在各个信息集中，轮到策略选择的参与者必须具有一个“推断”(Belief)值来确定博弈到达信息集中各个节点的可能性。对多节点信息集，“推断”值就是到达信息集中各个节点的概率分布，而对单节点信息集，则“推断”值对应的概率为1；②给定各参与者的“推断”值，则选择的策略应是“序列理性”(Sequentially Rational)的，也就是说，给定轮到策略选择的参与者的“推断”值，则该参与者在接下来的策略选择中必须使自己的效用最大；③若信息集在均衡路径上，则“推断”值由各参与者的均衡策略和贝叶斯法则共同确定；④若信息集不在均衡路径上，则“推断”值由各参与者可能有的均衡策略和贝叶斯法则共同确定[19]。

上述涉及的纳什均衡、“子博弈完美纳什均衡”和“完美贝叶斯均衡”具有内在联系。“子博弈完美纳什均衡”是“完美贝叶斯均衡”的特例，也就是说，“完美贝叶斯均衡”在完全且完美信息动态博弈中就是“子博弈完美纳什均衡”[19]。而在静态博弈中，完美贝叶斯均衡就是纳什均衡。

2.2.6 不完全信息动态博弈

不完全信息动态博弈(Incomplete Information Dynamic Game)又称“动态贝叶斯博弈”(Dynamic Bayesian Game)。与不完全信息静态博弈相比，不完全信息动态博弈中的博弈有

时间先后顺序，后参与者可以通过观察先参与者的行动，获得有关先参与者的信息，从而修正或证实自己对先参与者的策略。与不完全信息静态博弈类似，通过海萨尼转换方法，不完全信息动态博弈可以转变为完全但不完美信息动态博弈[19]。

在不完全信息动态博弈中，首先，"自然"选择参与者的类型，并将类型告诉参与者自己，但不告诉其他参与者，只将类型分布告诉其他参与者；在"自然"选择之后，参与者开始行动并有先后顺序，后行动者能观察到先行动者的行动，而不能观察到先行动者的类型[16]。但是，因为参与者的行动依赖于类型，每个参与者的行动都向后行动者传递着有关自己类型的某种信息，后行动者可以通过观察先行动者所选择的行动来推断先行动者的类型或修正对先行动者类型的"先验推断"(Prior Belief)，其实质是一种概率分布，然后，根据这一"推断"值选择自己的最优行动[16]。然而，先行动者并不是消极地选择行动，他预测到自己的行动将被后行动者所利用，就会设法选择传递对自己最有利的信息，避免传递对自己不利的信息[16]。这样，博弈过程不仅是参与者选择行动的过程，还是参与者不断调整"推断"值的过程。

由于不完全信息动态博弈通过海萨尼转换方法可以转变为完全但不完美信息动态博弈，因此与不完全信息动态博弈相关的均衡概念也是"完美贝叶斯均衡"，它汲取了"子博弈完美纳什均衡"和"贝叶斯均衡"的精华，是"贝叶斯均衡"、"子博弈完美均衡"和"贝叶斯推断"的结合[16]。

声明博弈是一类特殊的不完全信息动态博弈模型，这种博弈模型主要研究在有私人信息、信息不对称的情况下，人们采用口头或书面的声明来传递信息的博弈问题[15]。信号博弈是一种一般的具有信息传递机制作用的不完全信息动态博弈模型，它的基本特征是博弈方分为信息发出方和信号接收方两类，先行动的信号发出方的行为对后行动的信号接收方来说，具有传递信息的作用[25, 26]。

2.2.7　合作博弈

合作博弈(Cooperation Game)和非合作博弈是博弈论中最基本的一种分类，它们主要根据参与者的行为逻辑差别进行区分。一般地，将允许存在约束力协议的博弈称为合作博弈，而不存在有约束力协议的博弈称为非合作博弈[24]。前面介绍的博弈类型都属于非合作博弈的范畴。

事实上，合作博弈中存在有约束力的协议，这说明了参与合作博弈中的参与者之间存在共同利益，但这些利益又不完全一致。因为如果参与者之间利益完全一致或完全对立，就不需要协调或没有协调的余地，那就可以用个体理性决策(即通过非合作博弈)解决问题，那样就不再需要什么协议。因此，只有在参与者之间既存在不完全一致但又有共同利益的情况下，才可能需要利用协议来约束行动以实现更大的自身和共同利益[17]。由于利益不完全一致，又进一步决定了利益的分配，并促进善于利益分割的讨价还价(Bargain)的形成。实际上，合作博弈协议的内容除了利益分配以外就是约定具体的行动，而要达成协议的前提就是通过讨价还价在利益分配方面达成一致[17]。因此，不管合作博弈问题来源于经济交易、合作还是竞争，也不管参与博弈的人数多少，本质上都是关于利益分割的讨价还价[27]。

2.2.8　信号博弈

信号博弈(Signaling Game)实质是一种具有信息传递机制的不完全信息动态博弈。在

一个信号博弈中，有两个参与者 S 和 R，分别称为信号发送者(Sender，S)和信号接收者(Receiver，R)。他们在博弈时将先后选择自己的动作，其中参与者 S 的类型是私有信息，参与者 R 只有一个类型，且为共同信息。这就是说，参与者 R 具有不完全信息且参与者 R 可以从参与者 S 的行动中获得行为信息，这些行为信息对参与者 R 来说就是反映参与者 S 效用的信号。

由于信号博弈属于不完全信息动态博弈，因此可以通过海萨尼转换表示为完全但不完美信息动态博弈，其时间顺序如下：

(1)“自然”先按一定概率从参与者 S 的类型空间 Θ_S 中选择一个类型 θ_S，其中 $\theta_S \in \Theta_S$。参与者 S 知道 θ_S，但参与者 R 不知道。参与者 R 拥有对 θ_S的“推断”值(实质为先验概率)。

(2) 参与者 S 在观察到 θ_S后从其动作空间 A_S中选择一个动作 a_S，其中 $a_S \in A_S$。

(3) 参与者 R 观察到 a_S后，先应用贝叶斯法则从先验概率得到后验概率(即下一个“推断”值)，再从其动作空间 A_R中选择一个动作 a_R，其中 $a_R \in A_R$。

(4) 双方支付分别由 $u_S(\theta_S, a_S, a_R)$和 $u_R(\theta_S, a_S, a_R)$给出。

与信号博弈相关的均衡是完美贝叶斯均衡，包括纯策略或混合策略完美贝叶斯均衡。需要注意的是，“阶段博弈”实质是一种不完全信息静态博弈，因此其相关的均衡是纯策略或混合策略贝叶斯均衡。

一个信号博弈具有完美贝叶斯均衡的条件如下：

(1) 参与者 R 必须有关于参与者 S 类型的“推断”值，由于该“推断”值是在观察到 a_S之后作出的，因此记为 $p(\theta_S \mid a_S)$并满足

$$\forall \theta_S, p(\theta_S \mid a_S) \geqslant 0 \quad 且 \quad \sum_{\theta_S} p(\theta_S \mid a_S) = 1 \tag{2-1}$$

(2) 给定推断 $p(\theta_S \mid a_S)$和参与者 S 发出的信号 a_S，参与者 R 选择的行动 a_R^* 应该是最优的，也就是最优化问题，即

$$\max_{a_R} \sum_{\theta_S} p(\theta_S \mid a_S) u_R(\theta_S, a_S, a_R) \tag{2-2}$$

的解。

(3) 给定参与者 R 的最优行动 α_R^*，参与者 S 选择的动作 a_S^* 应该是最优的，也就是最优化问题，即

$$\max_{a_S} u_S(\theta_S, a_S, a_R) \tag{2-3}$$

的解。

(4) 对每个 $a_S \in A_S$，如果 $\exists \theta_S \in \Theta_S$ 使得 $a_S^* = a_S$，那么在对应 a_S 的参与者 R 的信息集中，参与者 R 的下一个“推断”值由贝叶斯法则得到，即

$$p(\theta_S \mid a_S) = \frac{p(\theta_S)}{\sum_{\theta_S} p(\theta_S)} \tag{2-4}$$

2.2.9 演化博弈

传统博弈类型(包含合作博弈和非合作博弈)假定参与者的博弈过程具有完全理性(Full Rationality)，也就是说，参与者在复杂的博弈环境中，对于博弈时相互的动作、支付等信息有准确的理解、分析和判断能力，已充分了解并遵守博弈规则，通过复杂且多层次的交

互推理得到博弈的结果——均衡。在这个过程中，参与者不会犯错误，不会怀疑对方的推理能力和理性，能准确地进行推理。

与传统博弈类型不同，演化博弈(Evolutionary Game)假定博弈的参与者在具有有限理性(Bounded Rationality)的基础上，分析参与者进行的策略选择，得到的是有限理性下的博弈均衡。这里的有限理性代表了参与者有一定的统计分析能力和对不同策略下得到收益的事后判断能力，但缺乏事前的预测和判断能力[18]。参与者只有有限的认知水平、有限的信息收集能力及有限的信息处理和推理能力，参与者的决策行为将受到其所处的群体环境的影响，只能通过学习、模仿进行策略选择。正是因为存在有限理性，参与者在演化博弈中不会马上得到最优的策略，而是需要在所处环境的影响下经历一个自我适应的调整过程，通过不断的学习、不断的试错找到最优的策略。这意味着演化博弈中的均衡不是一次选择的结果，而是需要动态地调整和适应才能达到，并且即使达到了均衡，在环境改变的前提下，可能会出现偏离现象。

演化稳定策略是演化博弈中的重要概念，其实质是演化博弈中的均衡，它源于生物进化论中的自然选择原理[18]。若一个种群达到了演化稳定策略，那么该种群中所有个体都采取这种策略，即使出现突变策略也不会影响到这个种群。也就是说，那些具备有限理性的种群个体根据其当前收益会不断地进行策略调整以实现其收益的最优化，最终达到一种动态平衡状态(即每个个体都选择演化稳定策略)。当一个种群达到演化稳定策略后，任何一个个体都不会单方面改变其策略，因为这种改变势必会造成个体收益的减少。所以，一个种群具有演化稳定策略就意味着该种群具有很大的稳定性，它将能抑制任何变异对种群的干扰。

演化稳定策略具有以下的重要性质：

(1) 演化稳定策略是一种对称的、完美的均衡[19]。

(2) 演化稳定策略代表了静态概念，在多种情况下可以直接从博弈模型的支付矩阵中得到演化稳定策略[19]。

(3) 纳什均衡不一定是演化稳定策略，只有达到严格纳什均衡才一定是演化稳定策略；反过来，演化稳定策略肯定是纳什均衡，其实质是纳什均衡的精炼[19]。

(4) 若一个对称的策略组合是纳什均衡，那么它是演化稳定策略[19]。

实际上，演化博弈的过程归根结底建立在选择(Selection)和突变(Mutation)这两大机制上。选择机制是指当前能够获得较高适应度(Fitness)的策略在今后会被更多的参与者通过学习模仿后采用；突变机制是指种群中的部分个体以随机的方式选择动作策略，这种突变可能会使参与者获得较高收益也可能获得较低收益，其中获得较高收益的策略经过选择机制的作用变得广泛流行，而获得较低收益的策略则自然消亡[19]。若将这种突变机制体现到种群的个体数量上，则采取广泛流行策略的个体数量将增加，而采取自然消亡策略的个体数量将减少。所以说，演化博弈的基本思想就是不断地演进、不断地自适应调整，从而使有较高收益的策略变得更加流行，直至达到演化稳定策略。

复制动态模型是目前描述种群个体行为选择机制的一种典型动力学模型，这是一种确定性和非线性模型。通过复制动态模型，可以较好地体现种群个体行为的有效理性变化趋势，在此基础上加入种群个体的随机选择策略行为后，就构成了一个包含选择和变异这两大机制的演化博弈模型，由此推出的结论能较好地预测种群个体的策略选择趋势[19]。复制动态动力学方程的给出主要基于使用某一策略的个体的增长率等于使用该策略时个体所得的

收益与种群平均收益的差[19]。下面给出复制动态动力学方程的表达形式。

设

$$S = \{s_1, s_2, \cdots, s_k\} \tag{2-5}$$

为一个种群中各个体可选择的动作组成的纯策略空间；$\phi_i(t)$为种群个体在时刻 t 选择纯策略 s_i的数量；

$$\theta(t) = \{\theta_1(t), \theta_2(t), \cdots, \theta_k(t)\} \tag{2-6}$$

为整个种群在时刻 t 所处的状态，该状态实际上可理解为该种群在时刻 t 的混合策略，其中，$\theta_i(t)$为种群个体在时刻 t 选择纯策略 s_i的数量占整个种群的比例，即

$$\theta_i(t) = \frac{\phi_i(t)}{\sum_i \phi_i(t)} \tag{2-7}$$

其中 $\theta_i(t)$满足

$$\sum_i \theta_i(t) = 1 \tag{2-8}$$

$u(s_i, \theta(t))$为种群个体选择纯策略 s_i的期望收益，即

$$u(s_i, \theta(t)) = \sum_j \theta_j(t) u(s_i, s_j) \tag{2-9}$$

$\bar{u}(\theta(t), \theta(t))$为整个种群的平均期望收益，即

$$\bar{u}(\theta(t), \theta(t)) = \sum_i \theta_i(t) u(s_i, \theta(t)) \tag{2-10}$$

假设每个个体的繁殖率与个体所占比例成正比[28]，即

$$\dot{\phi}_i(t) = \phi_i(t) u(s_i, \theta(t)) \tag{2-11}$$

由此，可得到复制动态方程[28]为

$$\begin{aligned}\dot{\theta}_i(t) &= \frac{\dot{\phi}_i(t)\sum_i \phi_i(t) - \phi_i(t)\sum_i \dot{\phi}_i(t)}{\left(\sum_i \phi_i(t)\right)^2} \\ &= \frac{\frac{\dot{\phi}_i(t)}{\phi_i(t)}\sum_i \phi_i(t) - \sum_i \dot{\phi}_i(t)}{\sum_i \phi_i(t)} \cdot \frac{\phi_i(t)}{\sum_i \phi_i(t)} \\ &= \theta_i(t)\left(u(s_i, \theta(t)) - \bar{u}(\theta(t), \theta(t))\right)\end{aligned} \tag{2-12}$$

2.2.10 微分博弈

微分博弈(Differential Game)理论建立于 1965 年美国人 Rufus Isaacs 的 *Differential Games*[29]一书，该书是世界上第一部微分博弈专著，其出版标志着微分博弈的诞生，Isaacs 也因此被尊称为“微分博弈之父”。其主要内容是研究动态的追逃策略问题，描述的是由一位追捕者(Pursuer)和一位逃避者(Evader)所组成的零和微分博弈及其解法。在这个零和微分博弈中，追捕者的目标是获得最大化抓捕逃避者的策略，而逃避者的目标是获得最大化逃脱追捕者的策略，其中追捕者和逃避者的策略分别是各自的追捕和逃避路线[29]。由于逃避者的收益是追捕者的损失，反之也一样，所以这是一个零和微分博弈。1970 年，美国数学家 Avner Friedman 建立了微分博弈值与鞍点存在性理论[30-32]，奠定了微分博弈的数学理论

基础。随后，微分博弈理论的研究与应用有了很大的发展，定量与定性微分博弈、非合作与合作微分博弈、随机微分博弈、主从微分博弈等不同博弈类型问题的研究不断深入。在国内，张嗣瀛院士的《微分对策》[33]应该是最早的专著；2000 年，李登峰教授的《微分对策及其应用》[34]专著问世。这两本专著重点分析了微分博弈在军事、控制问题上的应用。

实际上，微分博弈将原来离散的博弈过程扩展到连续时间之上，也就是说，参与者可以在无限小的时间内改变各自的控制策略。因此，使用微分博弈可以描述连续动态博弈系统的演化过程。微分博弈理论类似于传统的最优控制理论，且使用类似的数学分析处理工具。不过最优控制理论主要考虑的是单个参与者为一个目标而进行的控制，而微分博弈则要考虑多个参与者对成本函数各自有不同的目标而分别进行的控制，且还要考虑参与者之间选择控制策略时的相互影响。由于在连续时间上描述参与者之间的最优策略相互关系往往比较困难，因此，在微分博弈理论中需要对各参与者的控制策略空间作出限制。其中较严格的限制即为“开环”(Open-loop)控制策略，该类型的控制策略要求参与者在博弈过程中得不到新的信息，所以，只能构造出一个随时间而变化的控制函数作为自己的控制策略，而不能根据参与者双方的实际博弈进程的观察来动态改变自己的控制策略[35]。比“开环”控制策略限制要弱一些的是“闭环”(Closed-loop)控制策略，该控制策略使参与者可以得到反馈信息，从而能动态实时地更改各自的控制策略。但为了能在数学上进行处理，一般假设其中的一个参与者不能直接观察到其他参与者的博弈变量，而只能观察到某种状态变量，另外，还需假设博弈过程具有马尔可夫性，即以往的博弈历史不会影响到后续的博弈过程，参与者仅根据当前状态变量的取值来决定自己应采取的控制策略[35]。

微分博弈的均衡解主要有开环纳什均衡(Open-loop Nash Equilibrium)、闭环纳什均衡(Closed-loop Nash Equilibrium)和反馈纳什均衡(Feedback Nash Equilibrium)[36]。

开环纳什均衡的解法有 3 个方面的特点：首先，在其他参与者都采用最优控制策略的条件下，每位参与者在选择最优控制策略时，不仅要考虑自己当前的瞬时成本，还要考虑博弈状态的变化进展对自己未来涉及的成本带来的影响；其次，博弈的最优状态取决于所有参与者的最优控制策略以及当前的时间点和状态，而在博弈开始时间的最优状态与博弈的开始状态相同；最后，在所有参与者都采用最优控制策略的条件下，而且参与者的这些最优控制策略只依赖于当前时间和开始状态的情况下，每位参与者的目标成本函数的变化取决于它在当前的瞬时成本、当前的状态和当前的目标成本函数等[19]。

与开环纳什均衡的解法类似，闭环纳什均衡的解法包含开环纳什均衡解法的前两方面的特点，但第三方面的特点有区别。在闭环纳什均衡解法中，每位参与者的目标成本函数的变化除取决于它在当前的瞬时成本和当前的目标成本函数外，还取决于状态的瞬时变化。

而反馈纳什均衡的解法包含两方面的特点。首先，当所有参与者都采用根据当前时间点和状态确定的最优控制策略时，参与者价值函数的值将随着时间的进展而转变，且在每一瞬间转变的减数等于它的瞬时成本，而状态的最优变化进展为价值函数值所带来的所有转变之和；其次，参与者的价值函数在最后时间点的值等于参与者在博弈结束后的终期成本[19]。

通常，在一个两人零和微分博弈中，参与者在逗留期$[0,T]$区间的目标成本函数为

$$J(\mu(t),\nu(t)) = \int_0^T g(t,\boldsymbol{x}(t),\mu(t),\nu(t))\mathrm{d}t + q(\boldsymbol{x}(T)) \tag{2-13}$$

式中，$t\in[0,T]$为博弈的每一时刻；T 为博弈的结束时间；$\mu(t)$和 $\nu(t)$分别为两个参与者可以在时刻 t 采取的控制策略，博弈过程中使用的所有控制策略的集合代表了参与者随时间而进展的控制策略路径；$\boldsymbol{x}(t)$为状态向量，其动态变化过程常使用微分式

$$\begin{cases}\dot{\boldsymbol{x}}(t)=f(t,x(t),\mu(t),\nu(t))\\ \boldsymbol{x}(0)=\boldsymbol{x}_0\end{cases}\tag{2-14}$$

描述；$g(t,\boldsymbol{x}(t),\mu(t),\nu(t))$为参与者在时刻 t 的瞬时成本；$q(\boldsymbol{x}(T))$为博弈的终期成本。选择控制策略 $\mu(t)$的参与者在接下来的博弈过程中将试图最小化目标成本函数 $J(\mu(t),\nu(t))$，与之相反，选择控制策略 $\nu(t)$的参与者将试图最大化 $J(\mu(t),\nu(t))$。尤其需要说明的是，零和微分博弈的鞍点(Saddle-point)即是该微分博弈的纳什均衡，也就是说，在两个参与者都采用鞍点控制策略时，在对方没有改变控制策略的前提下，任何一方都不会偏离鞍点控制策略。因此，鞍点控制策略实际上已成为参与者能够选择的最优控制策略。

2.2.11 随机博弈

随机博弈(Stochastic Game)是一类具有状态概率转移的动态博弈，它由一系列阶段组成[15]。在随机博弈中每一"阶段博弈"的起始，博弈处于某种特定状态。每个参与者选择某种动作策略，此时会获得由当前状态和动作策略确定的收益。然后整个随机博弈按照概率的分布和参与者选择的动作策略随机转移到下一个"阶段博弈"。在新的"阶段博弈"(状态)，重复上一次的动作策略选择过程，再继续进行有限或无限次数的"阶段博弈"。最后，一个参与者得到的累积收益常用各"阶段博弈"的收益的贴现和或是各"阶段博弈"的收益的平均值的下限来计算。

如果整个随机博弈具有有限数量的参与者并且每个"阶段博弈"包含的状态数量有限，那么该随机博弈存在一个纳什均衡[15]。同样地，对于一个具有无穷阶段的随机博弈，如果使用各"阶段博弈"的收益的贴现和来计算参与者在整个随机博弈的收益，那么这个随机博弈也存在纳什均衡。Nicolas Vieille 已经证明具有有限阶段和有限状态的两人随机博弈当中，如果参与者在博弈过程中的收益使用各个阶段收益平均值的下限来计算，是能逼近纳什均衡的[15]。然而，包含两个以上的参与者的随机博弈是否存在纳什均衡，仍然是个未决的问题[15]。

下面给出双人零和随机博弈的形式化描述。在一个双人零和随机博弈中，设包含 z 个"阶段博弈"$\boldsymbol{\Gamma}_k(k=1,\ \cdots,\ z)$。每一个"阶段博弈"

$$\boldsymbol{\Gamma}_k=(\mu_{ij}^k)\tag{2-15}$$

是一个 $m_k\times n_k$ 矩阵，其每个矩阵元素为

$$\mu_{ij}^k=r_{ij}^k+\sum_{l=1}^{z}q_{ij}^{kl}\boldsymbol{\Gamma}_l\tag{2-16}$$

其中，对

$$\forall k,i,j,q_{ij}^{kl}\geqslant 0\quad 且\quad \sum_{l=1}^{z}q_{ij}^{kl}<1\tag{2-17}$$

当整个随机博弈结束时，"阶段博弈"$\boldsymbol{\Gamma}_k$ 的转移概率为

$$q_{ij}^{k0}=1-\sum_{l=1}^{z}q_{ij}^{kl}\tag{2-18}$$

参与者 1 的混合策略$\boldsymbol{\alpha}^k$ 是一个 m_k维向量并满足

$$\sum_{i=1}^{m_k} \alpha_i^k = 1 \tag{2-19}$$

其中，$\alpha_i^k \geqslant 0$。参与者 2 的混合策略$\boldsymbol{\beta}^k$ 是一个 n_k维向量并满足

$$\sum_{j=1}^{n_k} \boldsymbol{\beta}_j^k = 1 \tag{2-20}$$

其中，$\boldsymbol{\beta}_j^k \geqslant 0$。

给定参与者 1 和 2 的"策略对"(i,j)，可以计算从"阶段博弈"$\boldsymbol{\Gamma}_k$ 开始的期望支付$\nu_k(k=1,\cdots,z)$，从而可得到"策略对"(i,j)的博弈值向量$\boldsymbol{v}=(v_1,v_2,\cdots,v_z)$。如果博弈值向量$\boldsymbol{v}$存在，为计算参与者 1 和 2 的最优策略，需要将"阶段博弈"$\boldsymbol{\Gamma}_k$ 用期望支付

$$\nu_k = \mathrm{val}(\boldsymbol{\Delta}_k) \tag{2-21}$$

代替，其中 $\mathrm{val}(\boldsymbol{\Delta}_k)$是矩阵博弈$\boldsymbol{\Delta}_k$ 的值，且

$$\boldsymbol{\Delta}_k = (\nu_{ij}^k) \tag{2-22}$$

是一个 $m_k \times n_k$ 矩阵，其每个矩阵元素

$$\nu_{ij}^k = r_{ij}^k + \sum_{l=1}^{z} q_{ij}^{kl} \nu_l \tag{2-23}$$

最终，对整个随机博弈而言，参与者 1 和 2 的最优策略即是每个矩阵博弈$\boldsymbol{\Delta}_k$ 中所有各自最优策略的集合。

2.2.12 联盟博弈

联盟博弈(Coalitional Game)在合作博弈领域是应用最广泛的博弈[37-40]。联盟博弈使用联盟式描述，包含参与者集合和特征函数(Characteristic Function)两个元素。与联盟博弈相关的重要概念主要有"优超"核(Core)、夏普里值(Shapley Value)和稳定集(Stable Set)，其中稳定集是联盟博弈的解概念。

联盟博弈的最大优势在于所有参与者的收益都会有一定程度增加，或者至少有一个参与者的收益会在参与者相互的合作中有所增加，而其他参与者的收益都不会因此减小，因此一个联盟的整体收益会相应增加。实质上，这种收益的增加主要是因为联盟博弈选择的是合作行为，或者说是相互妥协的方式，这样就可以产生超出各个参与者单独采取博弈行为所获得的收益之和。当然，其实现过程需要参与联盟博弈的各个参与者在合作之前通过重复的讨价还价才能达成合作的共识。

联盟博弈的存在需要满足以下两个条件：

(1) 从联盟外部来看，联盟的整体收益要大于各个联盟内部参与者在非合作博弈中的收益的总和。

(2) 从联盟内部来看，应具有包含帕累托改进特性的分配规则，即每个合作参与者都能够获得一部分多于其不选择加入联盟时的收益。

在实际应用中，联盟博弈主要用来描述一群参与者之间合作的动态过程，处理合作群体的形成问题，使用 merge-and-split 规则动态更新联盟集合，协调参与者之间的行动，使得整个联盟的效用最大，个体参与者的收益最优。联盟博弈的 merge-and-split 规则能够以分布式的方式实现，适用于无线网络节点之间相互合作的博弈，它为无线网络设计公平的、健壮的、高效的合作通信策略提供了强有力的数学工具。

联盟博弈主要由参与者集合 $N=\{1,\cdots,n\}$ 和联盟值组成，其中，联盟值通常用 v 表示，它代表博弈中整个联盟的效用；联盟博弈表示为 (N,v)。联盟博弈具有可传递性(TU)和不可传递性(NTU)，可传递性是指联盟接收的总效用能以任何方式在联盟成员中分配。联盟博弈具有以下的基本定义。

定义 2-1 当联盟中仍有参与者加入或退出发生时，联盟博弈处于不稳定状态。当参与者没有动机形成新的联盟时，联盟博弈处于稳定状态，此时的稳定联盟叫做具有 TU 联盟核，可表示为

$$C_{\mathrm{TU}}=\left\{\theta:\sum_{i\in N}\theta_i=v(\theta)\quad 且\quad \forall S\subseteq N,\quad \sum_{i\in S}\theta_i\geqslant v(S)\right\} \tag{2-24}$$

定义 2-2 如果每个参与者获得的收益不小于单独行动时获得的收益，即

$$\forall i,\quad \theta_i\geqslant v(\{i\})\quad 且\quad \sum_{i\in N}\theta_i=v(N) \tag{2-25}$$

则收益向量 $\boldsymbol{\theta}=(\theta_1,\cdots,\theta_M)$ 反映出个体参与者是理性的。

定义 2-3 具有可传递性(TU)的联盟博弈 (N,v)，如果对于任何两个不相交的联盟，$S_1,S_2\subset N,S_1\cap S_2=\varnothing$，满足 $v(S_1\cup S_2)\geqslant v(S_1)+v(S_2)$，则此联盟具有超可加性。

定义 2-4 如果联盟的 TU 核为空或者很大且无法选择适当的收益分配集合时，则对于每个参与者 $i\in N$，由 Shaplay 值分配的收益为

$$\phi_i(v)=\sum_{S\subseteq N\setminus\{i\}}\frac{|S|!(|N|-|S|-1)!}{|N|!}[v(S\cup\{i\})-v(S)] \tag{2-26}$$

定义 2-5 对于给定的两个联盟集合 $S=\{S_1,\cdots,S_l\}$ 和 $R=\{R_1,\cdots,R_p\}$，定义 $\triangleright_i$ 为联盟集合 S 和 R 传递的二元关系，$S\triangleright_i R$ 表示联盟博弈参与者 i 偏向于加入联盟集合 S。

2.3 小结

本章概要介绍了博弈论的相关知识。通过对博弈论中重要的基本概念简介，对完全信息静态博弈、完全且完美信息动态博弈、重复博弈、不完全信息静态博弈、完全但不完美信息动态博弈、不完全信息动态博弈、合作博弈、信号博弈、演化博弈、微分博弈、随机博弈、联盟博弈等博弈类型的说明和分析，读者可以初步了解博弈论，为后续章节博弈论的应用和相关工作的比较提供知识准备。实际上，后续章节涉及的博弈类型要么是上述类型的一种，要么是上述类型的进一步延伸。

第 3 章

基于信号博弈的无线传感器网络入侵检测最优策略研究

本章应用信号博弈描述并分析恶意传感器节点和无线传感器网络入侵检测系统之间的交互过程。在每个独立的阶段，建立"阶段入侵检测博弈"模型，分别得到该模型的纯策略贝叶斯均衡和混合策略贝叶斯均衡。随着博弈的连续进行，构建"多阶段动态入侵检测博弈"反映恶意传感器节点和入侵检测代理之间的交互活动，得到相应的完美贝叶斯均衡，在此基础上实现入侵检测启动最优策略的机制和算法，从而在保证无线传感器网络安全的前提下，实现无线传感器网络入侵检测系统的节能目标。

3.1 引言

近年来，大量无线传感器网络基础理论和关键技术的研究为其大范围的应用奠定了基础，这些应用可包括人们日常生活的诸多方面。Akyildiz 等人[2]将无线传感器网络的应用领域分为军事、环境、健康、家庭以及其他商业领域。可以说，在不久的将来，由大量低成本传感器节点组成的无线传感器网络将成为人们生活中必不可少的一部分。

类似于其他计算机网络环境，无线传感器网络安全虽然不是必需的功能，但提供这种安全机制是保证无线传感器网络可用和可靠的必然需求。当前，保证网络安全的机制主要包括预防(Prevention)和检测(Detection)两大机制。预防机制主要通过加解密、密钥管理、安全路由、安全数据融合等技术提供数据的机密性(Confidentiality)、完整性(Integrity)和真实性(Authentication)；而检测机制通常使用入侵检测系统(Intrusion Detection System)实现，典型的有 DTRAB[41]。由于一个无线传感器网络可能包含大量恶意传感器节点，仅使用作为第一防线的预防机制不足以保证无线传感器网络的安全。这些恶意传感器节点的目的是在最小化被捕获概率的前提下最大化地破坏无线传感器网络的通信数据等，通过干扰无线传感器网络的正常工作和浪费正常传感器节点的珍贵资源获得利益。为了减少这些恶意节点造成的影响，无线传感器网络需要入侵检测系统检测那些已突破第一道防线的恶意传感器节点。通过使用入侵检测系统，使得无线传感器网络具备响应和隔离入侵者的能力，从而保证无线传感器网络的正常工作。

然而，要在无线传感器网络中资源有限的传感器节点上有效使用入侵检测系统，一个首先要解决的问题是如何选择合适的检测策略，因为这决定了无线传感器网络资源被消耗的

程度。无线传感器网络中的传感器节点在计算能力、存储容量和通信带宽等方面与现存网络相比有很大的不足。尤其是目前大多数传感器节点采用电池供电，因此能量较少。另外，应用入侵检测系统本身就需要耗费较多的计算和能量资源，这对于传感器节点而言是一个很大的负担。虽然伴随着微电子技术、计算机网络通信的发展，无线传感器网络的计算等能力将逐步提升，但如何节省传感器节点资源的消耗始终是一个需要考虑的问题。

博弈论作为一种研究参与者之间竞争和合作关系的数学理论工具[21]，已广泛应用于网络安全领域，如 P2P 安全[42]、防御 DOS 攻击[43-47]和入侵检测[48-53]等。博弈论包含适合于不同状况的博弈类型，如果要研究的问题中参与者具有不完全信息且博弈过程具有多个阶段时，那么信号博弈是一种非常合适的博弈类型，因为这种博弈模型具有“推断”（先验概率或后验概率）值动态更新的能力。

无线传感器网络入侵检测问题中的恶意传感器节点和入侵检测系统之间的交互可以方便地应用信号博弈进行描述。一个无线传感器网络入侵检测系统通常包含监测和决策模块。监测模块主要用于检查无线传感器网络中的被监控数据信息，而决策模块用于判断这些被监控的数据信息是否合法。其中，被监控数据信息包含了正常和恶意传感器节点所有的行为信息。因此，从检查到决策的整个动态过程实际上就是恶意传感器节点和入侵检测系统进行交互的过程。通过信号博弈这种数学工具，能很好地描述出这种交互过程的本质，实现入侵检测系统优化防御策略的要求，达到改进入侵检测系统正收益和有效节约传感器节点能量消耗的目的。

本章将应用信号博弈描述并分析恶意传感器节点和无线传感器网络入侵检测系统之间的交互过程。其中，无线传感器网络使用分布—集中混合式（Distributed-centralized）网络结构模型且每个传感器节点上已安装入侵检测代理（Intrusion Detection Agent），这些入侵检测代理构成了整个无线传感器网络入侵检测系统。为了节省能量消耗和减少数据包碰撞，不是所有的入侵检测代理而是仅位于簇头上的入侵检测代理才可能被启动实现对恶意传感器节点的入侵检测。在每个独立的阶段，建立“阶段入侵检测博弈”（Stage Intrusion Detection Game）模型，探索该博弈模型纳什均衡存在的条件并将分别得到纯策略贝叶斯均衡和混合策略贝叶斯均衡。随着博弈的进行，通过构建“多阶段动态入侵检测博弈”（Multi-stage Dynamic Intrusion Detection Game）来反映恶意传感器节点和入侵检测代理之间的交互活动。在这个过程中，入侵检测代理将依据恶意传感器节点的行为动态地更新针对恶意传感器节点的“推断”值，从而相应地调整它的防御策略。另外，在得到“多阶段动态入侵检测博弈”的完美贝叶斯均衡的基础上，设计入侵检测运行机制和相应的算法。

在扩展作者前期工作[54]的基础上，本章工作主要包括以下内容：

（1）基于信号博弈构建一种“无线传感器网络入侵检测博弈”模型用于研究恶意传感器节点和入侵检测代理之间的策略选择，这个模型满足了入侵检测代理对传感器节点的类型（正常或恶意）未知的实际场景。

（2）建立并证明了“无线传感器网络入侵检测博弈”模型存在均衡的定理，这些定理为入侵检测代理在决定是否采取保卫（Defend）或空闲（Idle）策略时提供最优的策略，也就是说，使用这些最优策略将使入侵检测代理不必始终采取动作 Defend，这样可以节省因采取动作 Defend 导致的能量消耗。

（3）基于完美贝叶斯均衡设计无线传感器网络入侵检测系统运行机制和相应的算法。

（4）构建模拟实验验证“多阶段动态入侵检测博弈”模型的有效性。

本章其余章节安排如下：3.2 节综述相关工作并突出说明本章工作与其他相关工作的区别；3.3 节描述分布—集中混合式无线传感器网络入侵检测博弈模型，包括“阶段入侵检测博弈”模型及其纯策略贝叶斯均衡及混合策略贝叶斯均衡、“多阶段动态入侵检测博弈”模型及其混合策略完美贝叶斯均衡以及提出基于混合策略完美贝叶斯均衡的入侵检测机制并给出入侵检测算法；3.4 节通过实验分析“多阶段动态入侵检测模型”的特性；3.5 节给出本章小结。

本章用到的符号含义如下：

θ_S表示“成员传感器节点”(Member Sensor Node)S，如果 S 是“正常成员传感器节点”，则 $\theta_S=0$；否则 $\theta_S=1$。

θ_R表示无线传感器网络“簇头入侵检测代理”R。

$a_S(\theta_S=0)$表示“正常成员传感器节点”的动作。

$a_S(\theta_S=1)$表示“恶意成员传感器节点”的动作。

$A_S(\theta_S)$表示“成员传感器节点”的动作空间。

$a_R(\theta_R)$表示“簇头入侵检测代理”R 的动作。

$A_R(\theta_R)$表示“簇头入侵检测代理”R 的动作空间。

g_A表示“恶意成员传感器节点”的攻击收益。

g_C表示“正常/恶意成员传感器节点”的合作收益。

g_D表示“簇头入侵检测代理”R 采取动作 Defend 的收益。

c_A表示“恶意成员传感器节点”的攻击成本。

c_C表示“正常/恶意成员传感器节点”的合作成本。

c_D表示“簇头入侵检测代理”R 采取动作 Defend 的成本。

l_F表示“簇头入侵检测代理”R 的误报损失。

α 表示“簇头入侵检测代理”R 的检测率。

β 表示“簇头入侵检测代理”R 的误报率。

p 表示“成员传感器节点”是恶意节点的概率。

ρ 表示“恶意成员传感器节点”采取动作 Attack 的概率。

δ 表示“簇头入侵检测代理”R 采取动作 Defend 的概率。

ρ^* 表示“恶意成员传感器节点”采取动作 Attack 的均衡概率。

δ^* 表示“簇头入侵检测代理”R 采取动作 Defend 的均衡概率。

σ_S表示“恶意成员传感器节点”的策略。

σ_S^* 表示“恶意成员传感器节点”的均衡策略。

σ_R表示“簇头入侵检测代理”R 的策略。

σ_R^* 表示“簇头入侵检测代理”R 的均衡策略。

ρ_k表示“恶意成员传感器节点”在“阶段博弈”t_k采取动作 Attack 的概率。

ρ_k^* 表示“恶意成员传感器节点” 在“阶段博弈”t_k采取动作 Attack 的均衡概率。

δ_k表示“簇头入侵检测代理”R 在“阶段博弈”t_k采取动作 Defend 的概率。

δ_k^* 表示“簇头入侵检测代理”R 在“阶段博弈”t_k采取动作 Defend 的均衡概率。

σ_{S_k} 表示“恶意成员传感器节点”在“阶段博弈”t_k的策略。

$\sigma_{S_k}^*$ 表示“恶意成员传感器节点”在“阶段博弈”t_k的均衡策略。

σ_{R_k} 表示“簇头入侵检测代理”R 在“阶段博弈”t_k的策略。

$\sigma_{R_k}^*$ 表示“簇头入侵检测代理”R 在“阶段博弈”t_k的均衡策略。

3.2 相关工作

入侵检测作为一种积极主动的安全防护技术，提供了防范内部攻击、外部攻击的能力。入侵检测技术是无线传感器网络安全研究的重点与难点之一，已经得到国内外研究人员的积极关注并已有大量文献发表。在中国知网、ACM、IEEE Xplore、Engineering Village、ScienceDirect、Web of Science、SpringerLink 等数据库中能查到的有关无线传感器网络入侵检测的文献近 500 余篇。典型的方法和技术主要有流量预测[55, 56]、基于安全协议的入侵检测系统[57]、异常检测[58-60]、相似观测结果分组[61]、检测接收功率异常[62]、移动代理[63]、马尔可夫线性预测[64]、支持向量机[65-67]、误用检测[68]、协同防御[69]、组合粒子群优化和径向基函数[70]、生物免疫[71, 72]、判断接收信号强度值[73]、危险理论[74]、散列预测[75]、蚁群优化[76]、核 Fisher 判别分析[77]、局部联系对比搜索[78]、基于区域的节点欺骗检测[79]、计数器对称加密[80]等。

然而，运行入侵检测系统本身就需要较多的计算资源，面对无线传感器网络节点资源有限的现状，如何真正地将入侵检测系统应用到无线传感器网络是一个很有挑战性的问题。近年来，博弈论为入侵检测的研究提供了新颖的思路。将博弈论应用于入侵检测领域，可以在入侵者和入侵检测系统之间建立利益冲突的数学模型，在考虑有限资源的基础上权衡不同策略带来的开销，对入侵检测系统进行是否启动的决策，从而提高入侵检测系统的效率。

当前，已有多种博弈类型被用于包括无线传感器网络在内的不同网络环境下的入侵检测研究，主要包括非合作完全信息静态博弈[49-52, 81-83]、重复博弈[45, 53, 84, 85]、贝叶斯博弈[47, 86-91]，但研究何时启动入侵检测系统的文献并不多见。

Liang 和 Xiao[92]分别从非合作博弈和合作博弈角度综述了博弈论在网络安全中的应用。Manshaei 等人[93]分别在物理层安全、自组织网络安全、入侵检测系统、隐私保护、网络安全经济学、密码学 6 个领域综述了博弈论的应用。作者等人[94]综述了博弈论在无线传感器网络安全方面的应用，其中包括无线传感器网络入侵检测领域。Javidi 和 Aliahmadipour[95]综述了如何应用博弈论改善 Ad Hoc 网络中的入侵检测系统。

周四清等人[84]提出的无线传感器网络入侵检测重复博弈模型被用于检测和响应传感器节点的自私行为以加强网络节点的协作性能，利用传感器节点与其邻居节点进行的重复博弈过程，广播传感器节点的效用值，从而即时检测出无线传感器网络节点的自私行为。并引入惩戒机制惩罚无线传感器网络节点的自私行为，从而大大降低了传感器节点背离合作的可能性。李奕男等人[52]将非合作完全信息静态博弈引入到 Ad Hoc 网络的入侵检测系统中，建立了入侵检测博弈模型并得到了该模型的纳什均衡解。该模型能有效提高入侵检测率，降低误检率且网络开销较小。石进等人[81]利用非合作完全信息静态博弈处理入侵检测系统响应的收益及攻击者策略变化等问题，提出了一种动态入侵响应模型，得到了稳定、可靠的最优解。陈行和陶军[86]应用贝叶斯博弈研究无线网络中入侵检测参数调整问题，根据入侵检测博弈模型中的完美均衡设计了入侵检测时间间隔调整算法和参数修正算法，这

些算法有效地帮助无线网络入侵检测系统检测出发生变化的恶意攻击行为。王静等人[90]将基于贝叶斯博弈并结合节点激励机制的入侵检测模型运用于一种改进的安全路由协议，有效地抑制了节点的自私行为。严辉等人[85]利用重复博弈提出了一种适用于 Ad Hoc 网络的入侵攻击预测模型。通过建立入侵检测系统和入侵攻击者之间的博弈模型，计算阶段博弈的经典纳什均衡，并得到了重复博弈情况下的子博弈精炼纳什均衡，再使用最优反应均衡模型预测入侵攻击者和入侵检测系统在博弈阶段中选择不同策略的概率。赵柳榕等人[96]将博弈论用于建立虚拟专用网(VPN)和入侵检测系统的信息安全技术组合模型，从而为阻止黑客入侵和降低信息安全技术配置成本提供优化策略。Chen 和 Leneutre[51]利用非合作完全信息静态博弈建立拥塞攻击者和受攻击网络之间的模型，当达到纳什均衡时，实现通过增加拥塞攻击者的能量消耗促使其快速死亡的防范策略。在 Huang 等人[50]提出的马尔可夫 IDS(Markovian IDS)中，将非合作完全信息静态博弈与异常和误用检测相结合用于确定最佳的防护策略。Chen 和 Leneutre[82]将利用非合作完全信息静态博弈得到的理性攻击者和入侵检测系统的最优策略用于入侵检测系统的设计和部署中。Kantzavelou 和 Katsikas[53]利用重复博弈建立内部攻击和入侵检测系统之间的博弈模型，将一般的纳什均衡扩展到随机最优反应均衡(Quantal Responsse Equilibrium)来预测内部攻击者的行为。Zhu 等人[91]利用动态贝叶斯博弈建立了一个动态入侵检测自动响应系统，为入侵检测系统的配置提供了最优的配置方案。Rafsanjani 等人[87]将贝叶斯博弈用于确定何时启动入侵检测系统的阈值，一旦攻击概率超过该阈值，将启动节点上的 IDS 服务。Bedi 等人[97]将博弈论用于分布式拒绝服务攻击领域，为防御者在如何设置防火墙方面提供优化策略，以便有效阻止恶意数据流和保证正常数据流的通过。Shamshirband 等人[98]建立了包含 Sink 节点、基站、攻击者 3 个参与者的策略博弈模型，当无线传感器网络某节点流量超过限定的阈值时启动该模型，再利用合作博弈和模糊 Q-learning 算法为 Sink 节点和基站合作防御拒绝服务攻击提供了优化策略。Moosavi 和 Bui[99]利用"非零和不完全信息随机博弈"(Nonzero-sum Discounted Stochastic Games with Incomplete Information)分析无线传感器网络中的入侵检测问题，在参与者信息不确定的情况下给出了一种鲁棒的优化防御策略。而 Zonouz 等人[100]利用模糊逻辑理论(Fuzzy Logic Theory)分析网络级安全事件的基础上，采用"斯塔克尔伯格随机博弈"(Stackelberg Stochastic Game)得到了优化的入侵响应策略。

信号博弈在无线网络领域已有一些应用。刘玉枚等人[101]利用信号博弈解决 P2P 网络系统资源交易中存在的不完全信息问题，提出了一种资源定价机制。通过建立信号博弈模型模拟信息的不完全性，使得系统资源的需求者能区分所需资源质量，并通过引入一种调价机制实现资源价格调整。陈亚睿等人[102]针对云计算环境下如何确定不可信云终端用户并合理分析云用户的异常行为问题，提出了一种基于信号博弈的用户行为模型，在考虑入侵检测系统存在误报和漏报的情况下，利用"多阶段博弈"分析云终端用户的类型，结合用户的当前行动和历史行动，实现准确地推断云终端用户类型，为主动安全机制提供了理论基础。Patcha 和 Park[103]在基于主机的 Ad Hoc 网络入侵检测系统中利用信号博弈建立博弈模型，但未深入研究该模型的特性，如模型是否存在均衡点等。Wang 等人[104]研究无线传感器网络中恶意节点和正常节点之间的共存问题。实际上，即使一个恶意传感器节点已被准确检测，但也许它自身并不知道已被列入恶意节点，因此它可能通过伪装自己的方法表现出正常节点的功能。这样，这种恶意节点仍旧可以被保留并使用，从而为恶意节点和正常节点

都能带来收益。在这样的背景下，他们提出利用信号博弈建立恶意节点的检测模型，并得到了模型的纯策略和混合策略纳什均衡。随着博弈的持续，他们根据贝叶斯规则进行“推断”值的更新，并证明了这个动态的恶意节点检测博弈模型具有完美贝叶斯均衡。Estiri 和 Khademzadeh[105]针对无线传感器网络中的丢包攻击，利用信号博弈建立攻击者和节点之间的博弈模型，将攻击者和节点之间的交互关系通过不完全信息动态博弈进行描述，证明了该模型存在完美贝叶斯均衡，同时，说明了达到均衡点即得到了优化的防御策略。Li 等人[106]利用不完全信息动态博弈分析 Ad Hoc 网络正常节点和恶意节点之间的交互关系并建立了相应的博弈模型。其中正常节点根据对手的行为更新自己的“推断”值，给出是否报告恶意节点的理性决策。另一方面，恶意节点通过评估自己被捕获的风险来决定何时逃离以避免被惩罚的策略。Maia 等人[107]针对延迟容忍网络中大多数路由协议未考虑能量消耗的情况，利用信号博弈建立了多路数据转发模型，给出了基于目标节点累积能耗的路由优化策略。Paramasivan 等人[108]利用信号博弈分析正常和恶意节点的行为，通过使用完美贝叶斯均衡(Perfect Bayesian Equilibrium)策略，最小化了恶意节点的收益，促进了正常节点的相互合作。

与上述相关工作相比，本章采用信号博弈建立“成员传感器节点”与“簇头入侵检测代理”之间的博弈模型，并通过计算得到“阶段入侵检测博弈”的纯策略和混合策略贝叶斯均衡以及“多阶段动态入侵检测博弈”的混合策略完美贝叶斯均衡，这些均衡将为“簇头入侵检测代理”给出何时选择动作 Defend 的最优策略。本章思想部分来自文献[104]中的恶意传感器节点检测博弈模型，但在构建博弈的支付矩阵时，本章考虑了入侵检测系统的检测率和误报率，而他们[104]考虑的是通道的不可靠性和恶意传感器节点成功攻击的概率。因此，与文献[104]相比较，本章得到了不同的均衡结果。另外，本章内容集中于利用信号博弈决定无线传感器网络入侵检测系统何时启动的策略问题，而上述相关工作大都研究的是 Ad Hoc 网络环境。最后，本章使用与上述相关工作不同的网络模型，将入侵检测代理驻留在每个传感器节点上，但仅有簇头上的入侵检测代理根据信号博弈结果进行启动，这种网络模型非常有利于无线传感器网络的能量节省。

3.3 无线传感器网络入侵检测博弈模型

3.3.1 网络模型

根据入侵检测系统代理的安装位置，Farooqi 等人[109]将无线传感器网络入侵检测系统分为三类：纯分布式(Purely Distributed)、纯集中式(Purely Centralized)和分布—集中混合式(Distributed-centralized)。在纯分布式无线传感器网络入侵检测系统中，入侵检测代理被安装于每个传感器节点并在本地检查相邻传感器节点的恶意行为。而在纯集中式网络结构中，入侵检测代理被安装于基站(Base Station)上，这种结构常需要额外的路由协议用来收集传感器节点的数据信息并以此分析传感器节点的行为。由于采用聚簇结构的无线传感器网络具有能耗和控制负荷低的特点，为适应这种网络结构，分布—集中混合式被引入进来，这种方式将入侵检测代理仅安装在“监控传感器节点”(Monitor Sensor Node)上，而“监控传感器节点”除执行入侵检测外，还具有与正常节点一样的转发数据功能。

本章采用的无线传感器网络入侵检测系统网络结构属于分布—集中混合式。但与入侵检测代理仅被安装在“监控传感器节点”上的情况不同，在本章采用的网络结构中，所有传感器节点都已部署入侵检测代理。与此同时，因为采用聚簇(Clustering)技术能显著改善网络生存期[110]，所以本章将聚簇技术用于无线传感器网络以便形成相互连接的层次结构。通过采用这种聚簇技术，所有的传感器节点都被分配到不同的簇中。每个簇都有一个称为簇头(Cluster Head, CH)的协调者和一些“成员传感器节点”。所有的簇头形成层次结构中的高层节点，而所有的“成员传感器节点”组成了低层节点。在这样的层次结构中，“成员传感器节点”通过“责任簇头”(Responsible CH)发送数据，“责任簇头”汇聚数据并通过其他的簇头将数据传输到基站。为了平衡簇头传感器节点的能量消耗，该节点经常需要定期更新。与无线传感器网络中的平面结构(Flat Architecture)相比，这种聚簇结构在减少能量消耗和降低通道碰撞方面具有显著的优点。当一个能量充沛的传感器节点被选中作为一个簇头时，驻留在簇头上的入侵检测代理将被同时启动，而处于“成员传感器节点”上的入侵检测代理将处于休眠状态。因此，簇头除汇聚和发送数据外，还有入侵检测的功能。图 3-1 给出了本章的网络模型。

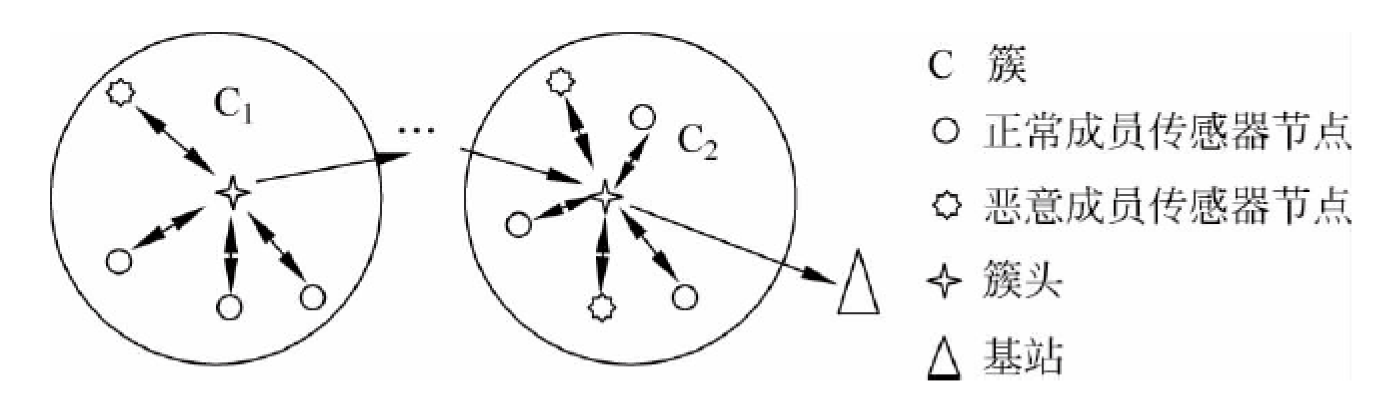


图 3-1　无线传感器网络入侵检测网络模型

在图 3-1 中，合法的传感器节点包括簇头和基站。“成员传感器节点”的类型可能是正常或恶意的。这些“成员传感器节点”知道它们自己的类型，但簇头不知道与它处于相同簇内的“成员传感器节点”类型。为适应这种网络环境，本章将采用信号博弈来描述“恶意成员传感器节点”的检测过程。当一个信号博弈被重复进行时，它可以被分为连续而独立的阶段，在每个阶段“成员传感器节点”和“簇头入侵检测代理”进行博弈的模型称为“阶段入侵检测博弈”。

3.3.2　阶段入侵检测博弈模型

定义 3-1　“阶段入侵检测博弈”是一个五元组$\mathbb{G}=(N,\Theta,A,P,U)$，其中：

- $N=\{$“成员传感器节点”S，“簇头入侵检测代理”$R\}$是一个包含两个参与者的集合。
- $\Theta=\Theta_S\times\Theta_R$，其中$\Theta_S=\{\theta_S=0,\theta_S=1\}$是“成员传感器节点”$S$的类型空间，$\Theta_R=\{\theta_R\}$是“簇头入侵检测代理”$R$的类型空间。
- $A=A_S\times A_R$，其中$A_S=\{A_S(\theta_S=0),A_S(\theta_S=1)\}=\{\{a_S(\theta_S=0)|\ \text{Cooperate}\},\{a_S(\theta_S=1)|\ \text{Attack, Cooperate}\}\}$是“成员传感器节点”$S$可用的动作集合，$A_R=\{a_R|\ \text{Defend, Idle}\}$是“簇头入侵检测代理”$R$可用的动作集合。
- $P:\Theta\mapsto[0,1]$是关于“成员传感器节点”S的推断(先验概率)，$P=(p,1-p)$，其中，p表示“恶意成员传感器节点”(Malicious Member Sensor Node)的概率，而$1-p$表示“正常成员传感器节点”(Normal Member Sensor Node)的概率。

• $U=\{(u_S,u_R)\}$，其中，$u_S:A\times\Theta\mapsto\mathbb{R}$ 是“成员传感器节点”S 的支付函数，$u_R:A\times\Theta\mapsto\mathbb{R}$ 是“簇头入侵检测代理”R 的支付函数，u_S 和 u_R 的支付值如表 3-1 所示。

表 3-1 “阶段入侵检测博弈”的支付矩阵

(a)“成员传感器节点”S 是恶意节点

a_R \ a_S	Defend	Idle
Attack	$(1-\alpha)g_A-\alpha g_D-c_A$，$\alpha g_D-(1-\alpha)g_A-c_D$	g_A-c_A，$-g_A$
Cooperate	g_C-c_C，$-\beta l_F-c_D$	g_C-c_C，0

(b)“成员传感器节点”S 是正常节点

a_R \ a_S	Defend	Idle
Cooperate	g_C-c_C，$-\beta l_F-c_D$	g_C-c_C，0

为反映无线传感器网络和入侵检测系统的特性，本章为“阶段入侵检测博弈”模型选择了一些特定的参数，当“恶意成员传感器节点”试图攻击无线传感器网络从而浪费其资源时，将影响无线传感器网络的正常运行，造成相邻节点通信的瘫痪，这个攻击过程将给“恶意成员传感器节点”带来收益，然而，它们也必须付出相应的成本用以支付它们的攻击。因此，对一个“恶意成员传感器节点”而言，本章引入 g_A 和 c_A 来分别表示攻击收益和成本。当一个“成员传感器节点”选择动作 Cooperate 时，意味着该节点能正常进行通信，也就是说，数据包能够被顺利地转发。这样，“正常成员传感器节点”将从具有良好通信保障的无线传感器网络中获取收益，而“恶意成员传感器节点”也将从它的伪装过程中获取收益。然而，在合作过程(即选择动作 Cooperate)中，接收和转发数据包都会消耗传感器节点的能量。为了简单起见，本章假设“恶意成员传感器节点”和“正常成员传感器节点”将得到相同的收益和付出相同的成本。因此，对一个“成员传感器节点”而言，本章引入 g_C 和 c_C 分别表示选择动作 Cooperate 的收益和成本。当“簇头入侵检测代理”选择动作 Defend 时，它将获得收益 g_D，这是因为它成功地检测到了“恶意成员传感器节点”。与此同时，“簇头入侵检测代理”必须付出相应的成本 c_D 用于支付能量的消耗。显然，与普通计算机网络中的入侵检测系统类似，“簇头入侵检测代理”中也存在检测率和误报率，本章分别用 α 和 β 表示。其中，存在误报率意味着“成员传感器节点”可能会在正常的通信中被误认为恶意节点，这对“簇头入侵检测代理”而言将造成损失 l_F。

在定义 3-1 给出的“阶段入侵检测博弈”模型中，总共有两个参与者，包括用 θ_S 表示的“成员传感器节点”S(发送者)和用 θ_R 表示的“簇头入侵检测代理”R(接收者)。“成员传感器节点”S 可能是正常的也可能是恶意的，分别用 $\theta_S=0$ 和 $\theta_S=1$ 表示，这些类型信息对“簇头入侵检测代理”R 而言都是私有信息。在每个阶段，每个参与者从它的动作空间中选择自己的动作，当“成员传感器节点”S 是恶意时，它可能采取攻击或合作行为，采取合作是因为它想伪装自己以避免被监测到，也就是说，传感器节点类型 $\theta_S=1$ 采取的动作 $a_S(\theta_S=1)$ 可能是 Attack 或者 Cooperate。而当“成员传感器节点”S 是正常节点时，它总是选择合作的行为，也就是说，类型 $\theta_S=0$ 的动作 $a_S(\theta_S=0)$ 总是 Cooperate。因此“成员传感器节点”S 的动作空间 $A_S(\theta_S)$ 可以表示为{Attack, Cooperate}。为了节省簇头节点的能量以便获得较长

的生存期，“簇头入侵检测代理”R 不应该总是选择动作 Defend，也就是说，有时它应该选择动作 Idle。这样传感器节点类型 θ_R 采取的动作 $a_R(\theta_R)$ 可能是 Defend 或 Idle。因此，“簇头入侵检测代理”R 的动作空间是{Defend，Idle}。表 3-1 给出了“阶段入侵检测博弈”的支付矩阵。

在表 3-1 中，除了动作 Idle 外，都将产生成本。对于“动作对”($a_S(\theta_S=1)=$Attack，$a_S(\theta_R)=$Defend)而言，传感器节点类型 $\theta_S=1$ 的支付等于未被检测到时的收益减去被检测到时的损失再减去攻击的成本，而传感器节点类型 θ_R 的支付等于成功检测到恶意节点的收益减去未检测到恶意节点的损失再减去检测的成本。对“动作对”($a_S(\theta_S=1)=$Attack，$a_S(\theta_R)=$Idle)而言，传感器节点类型 $\theta_S=1$ 的支付等于攻击获得的收益减去攻击的成本，而传感器节点类型 θ_R 的支付等于被恶意节点攻击造成的损失。对“动作对”($a_S(\theta_S=1)=$Cooperate，$a_R(\theta_R)=$Defend)而言，传感器节点类型 $\theta_S=1$ 的支付等于合作的收益减去合作的成本，而传感器节点类型 θ_R 的支付等于误报造成的损失减去采取动作 Defend 的成本。至于其他的“动作对”所产生的支付应该容易理解，在此不再赘述。

3.3.3 “阶段入侵检测博弈”的均衡

作为一种不完全信息动态博弈类型，“阶段入侵检测博弈”中的 “簇头入侵检测代理”R 虽然不知道“成员传感器节点”S 的类型，但这种博弈模型仍能得到纯策略和混合策略贝叶斯均衡。当然，要得到这些均衡，首先需要通过海萨尼转换将“阶段入侵检测博弈”转化为完全但不完美信息动态博弈。在转化时，根据不完全信息动态博弈的时间顺序，一个虚拟的参与者“自然”(Nature)被引入进来，这个“自然”将首先行动并以一定的概率确定“成员传感器节点”S 的类型。图 3-2 给出了转换后的“阶段入侵检测博弈”的扩展式。

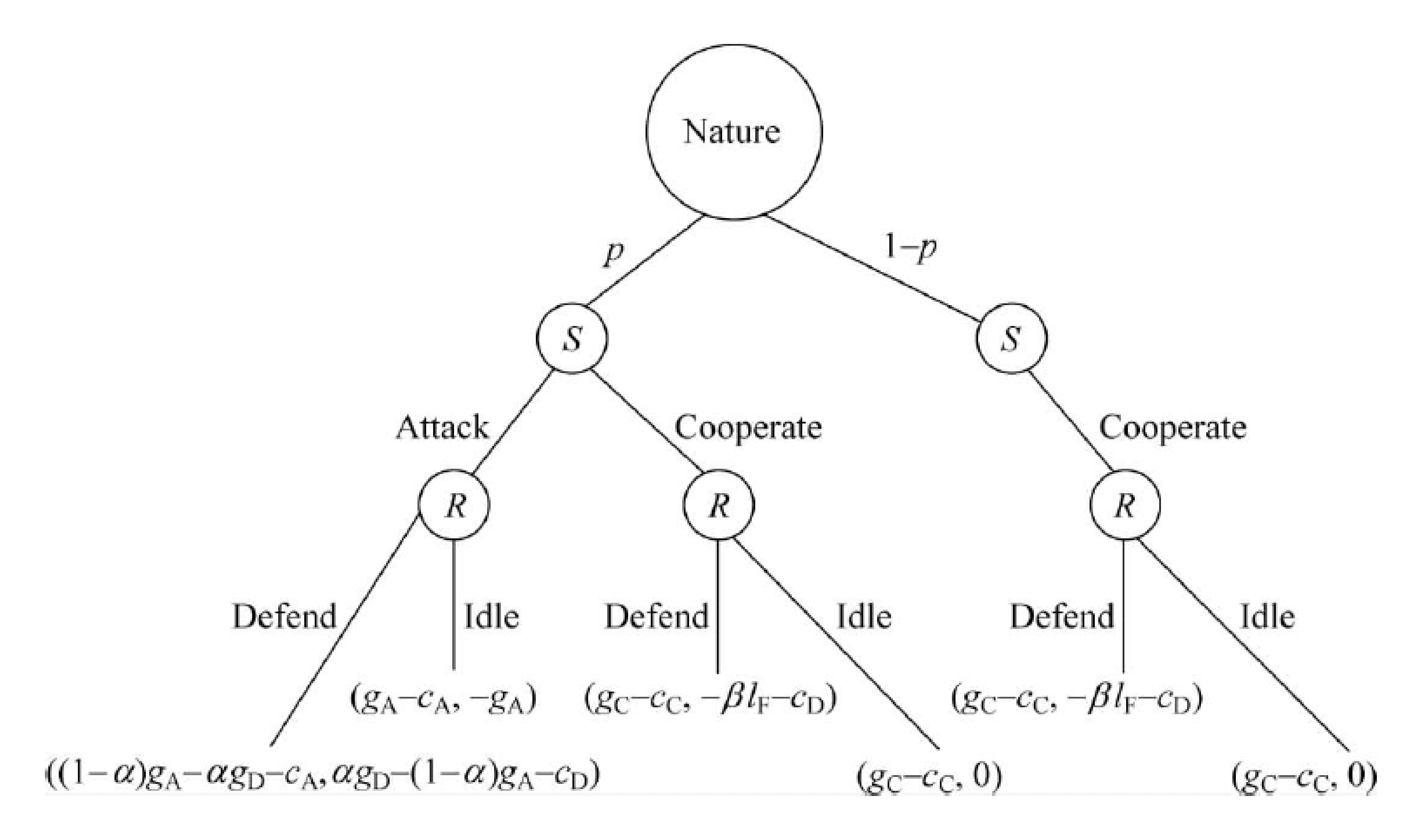


图 3-2 “阶段入侵检测博弈”的扩展式

定理 3-1 当 $p<(\beta l_F+c_D)/(\alpha g_D+\alpha g_A+\beta l_F)$成立时，“阶段入侵检测博弈”存在纯策略贝叶斯均衡。

证明 (1)“成员传感器节点”S 选择纯策略($a_S(\theta_S=1)=$Attack，$a_S(\theta_S=0)=$Cooperate)。这种情况意味着当“成员传感器节点”S 属于恶意节点时，总是选择动作 Attack，而属于正常节点时总是选择动作 Cooperate。因此，对“簇头入侵检测代理”R 而言，

选择 Defend 和 Idle 的期望收益分别是

$$Eu_R(\text{Defend}) = p(\alpha g_D - (1-\alpha)g_A - c_D) + (1-p)(-\beta l_F - c_D) \tag{3-1}$$

和

$$Eu_R(\text{Idle}) = -pg_A + (1-p)\cdot 0 = -pg_A \tag{3-2}$$

如果 $Eu_R(\text{Defend}) \geqslant Eu_R(\text{Idle})$，也就是说

$$p(\alpha g_D - (1-\alpha)g_A - c_D) + (1-p)(-\beta l_F - c_D) \geqslant -pg_A$$

即

$$p \geqslant (\beta l_F + c_D)/(\alpha g_D + \alpha g_A + \beta l_F) \tag{3-3}$$

那么"簇头入侵检测代理"R 的最优策略是采取动作 Defend。然而，当"簇头入侵检测代理"R 选择动作 Defend 时，Attack 将不再是"成员传感器节点"S 的最优动作，这是因为不等式

$$(1-\alpha)g_A - \alpha g_D - c_A < g_C - c_C \tag{3-4}$$

恒成立。因此，$\{(a_S(\theta_S=1)=\text{Attack},\ a_S(\theta_S=0)=\text{Cooperate}),\ a_R(\theta_R)=\text{Defend}\}$ 不是"阶段入侵检测博弈"的一个纯策略贝叶斯均衡。

如果 $Eu_R(\text{Defend}) < Eu_R(\text{Idle})$，也就是说

$$p < (\beta l_F + c_D)/(\alpha g_D + \alpha g_A + \beta l_F) \tag{3-5}$$

成立时，"簇头入侵检测代理"R 的最优策略是采取动作 Idle。相应地，Attack 将成为"成员传感器节点"S 的最优动作，这是因为不等式

$$g_A - c_A > (1-\alpha)g_A - \alpha g_D - c_A \tag{3-6}$$

恒成立。因此，$\{(a_S(\theta_S=1)=\text{Attack},\ a_S(\theta_S=0)=\text{Cooperate}),\ a_R(\theta_R)=\text{Idle}\}$ 是"阶段入侵检测博弈"的一个纯策略贝叶斯均衡。

(2) "成员传感器节点"S 选择纯策略 $(a_S(\theta_S=1)=\text{Cooperate},\ a_S(\theta_S=0)=\text{Cooperate})$。这种情况意味着不管"成员传感器节点"$S$ 是何种类型它总是选择动作 Cooperate。对"簇头入侵检测代理"R 而言，针对"成员传感器节点"S 的动作 Cooperate 的最优响应是选择动作 Idle，而对于"恶意成员传感器节点"$\theta_S=1$ 而言，针对"簇头入侵检测代理"R 的动作 Idle 的最优响应是选择动作 Attack。这样与纯策略 $(a_S(\theta_S=1)=\text{Cooperate},\ a_S(\theta_S=0)=\text{Cooperate})$ 相互矛盾，因此，$\{(a_S(\theta_S=1)=\text{Cooperate},\ a_S(\theta_S=0)=\text{Cooperate}),\ a_R(\theta_R)=\text{Idle}\}$ 不是"阶段入侵检测博弈"的一个纯策略贝叶斯均衡。

综上所述，当

$$p < (\beta l_F + c_D)/(\alpha g_D + \alpha g_A + \beta l_F) \tag{3-7}$$

成立时，"阶段入侵检测博弈"存在纯策略贝叶斯均衡 $((a_S(\theta_S=1)=\text{Attack},\ a_S(\theta_S=0)=\text{Cooperate}),\ a_R(\theta_R)=\text{Idle})$。这意味着"恶意成员传感器节点"总是会选择动作 Attack 并且"正常成员传感器节点"总是选择动作 Cooperate，而"簇头入侵检测代理"R 总是选择动作 Idle。证毕。

虽然"阶段入侵检测博弈"存在纯策略贝叶斯均衡，但这与实际的防御要求不符，因为根据定理 3-1，"簇头入侵检测代理"R 总是选择动作 Idle，这就是说，"恶意成员传感器节点"永远不会被捕获。因此仅得到纯策略贝叶斯均衡对"入侵检测博弈"是不够的，必须要找到能用于检测恶意传感器节点的混合策略贝叶斯均衡。

定理 3-2 当条件 $p \geqslant (\beta l_F + c_D)/(\alpha g_D + \alpha g_A + \beta l_F)$ 成立时，"阶段入侵检测博弈"存在混合策略贝叶斯均衡。

证明　显然，由定理 3-1，只有条件 $p \geq (\beta l_F + c_D)/(\alpha g_D + \alpha g_A + \beta l_F)$ 成立时，"阶段入侵检测博弈"才有可能存在混合策略贝叶斯均衡。

设"恶意成员传感器节点"$\theta_S = 1$ 的混合策略为 $\sigma_S = (\rho, 1-\rho)$，其中，$\rho$ 表示"恶意成员传感器节点"采取动作 Attack 的概率。那么，"簇头入侵检测代理"R 采取动作 Defend 和 Idle 的期望收益分别是

$$Eu_R(\text{Defend}) = \rho p(\alpha g_D - (1-\alpha)g_A - c_D) + (1-\rho)p(-\beta l_F - c_D) + (1-p)(-\beta l_F - c_D) \tag{3-8}$$

和

$$Eu_R(\text{Idle}) = \rho p(-g_A) + (1-\rho)\cdot p \cdot 0 + (1-p)\cdot 0 = -\rho p g_A \tag{3-9}$$

在"恶意成员传感器节点"$\theta_S = 1$ 采取最优混合策略 σ_S^* 前提下，由"簇头入侵检测代理"R 采取动作 Defend 和 Idle 的无差异性(Indifference)可以得到

$$Eu_R(\text{Defend}) = Eu_R(\text{Idle}) \tag{3-10}$$

因此，"簇头入侵检测代理"R 的最优混合策略为

$$\rho^* = \frac{\beta l_F + c_D}{p\alpha g_D + p\alpha g_A + p\beta l_F} \tag{3-11}$$

设"簇头入侵检测代理"R 的混合策略为 $\sigma_R = (\delta, 1-\delta)$，其中，$\delta$ 表示"簇头入侵检测代理"R 采取动作 Defend 的概率。那么，"成员传感器节点"采取动作 Attack 和 Cooperate 的期望收益分别是

$$Eu_S(\text{Attack}) = \delta p((1-\alpha)g_A - \alpha g_D - c_A) + (1-\delta)p(g_A - c_A) \tag{3-12}$$

和

$$Eu_S(\text{Cooperate}) = \delta p(g_C - c_C) + (1-\delta)p(g_C - c_C) + \delta(1-p)(g_C - c_C) + (1-\delta)(1-p)(g_C - c_C) \tag{3-13}$$

在"簇头入侵检测代理"R 采取最优混合策略 σ_R^* 的情况下，由"成员传感器节点"S 采取动作 Attack 和 Cooperate 的无差异性可以得到

$$Eu_S(\text{Attack}) = Eu_S(\text{Cooperate}) \tag{3-14}$$

因此，"成员传感器节点"S 的最优混合策略为

$$\delta^* = \frac{p g_A - p c_A - g_C + c_C}{p(\alpha g_A + \alpha g_D)} \tag{3-15}$$

综上所述，当条件 $p \geq \dfrac{\beta l_F + c_D}{\alpha g_D + \alpha g_A + \beta l_F}$ 成立时，"阶段入侵检测博弈"存在一个混合策略贝叶斯均衡($\sigma_S^*(a_S(\theta_S = 1) = \text{Attack}, a_S(\theta_S = 0) = \text{Cooperate})$，$\sigma_R^*(a_R(\theta_R) = \text{Defend})$)。"阶段入侵检测博弈"存在混合策略贝叶斯均衡意味着当"簇头入侵检测代理"R 以概率 δ^* 采取动作 Defend 时，"恶意成员传感器节点"将以概率 ρ^* 采取动作 Attack 而"正常成员传感器节点"总是选择动作 Cooperate。证毕。

根据定理 3-1 和定理 3-2，当"阶段入侵检测博弈"达到贝叶斯均衡时，"成员传感器节点"S 和"簇头入侵检测代理"R 在不同的概率 p 下都能选择它们各自不同的优化策略。从中可以看出，概率 p 实际上与"簇头入侵检测代理"的检测率 α 和误报率 β 有关。随着"推断"值的变大，由定理 3-2 中与概率 p、检测率 α 和误报率 β 相关的混合策略贝叶斯均衡可知，"恶意成员传感器节点"选择动作 Attack 的概率越来越小。使用上述"阶段入侵检测博

弈”的好处在于“簇头入侵检测代理”不必在每一个阶段选择动作 Defend，这样簇头用于执行入侵检测代理的能量消耗就变小了。接下来要面临的一个新问题是在每一个独立的阶段如何确定一个合理的“推断”值 p，这个“推断”值必须能根据实际的情况进行动态地更新。因此，本章根据无线传感器网络入侵检测的实际情况，进一步将“单阶段静态入侵检测博弈”扩展成“多阶段动态入侵检测博弈”，并且讨论其中的“推断”值是如何更新的。

3.3.4 多阶段动态入侵检测博弈模型

随着“成员传感器节点”S 和“簇头入侵检测代理”R 交互的进行，在每一个连续的阶段 $t_k(k=1,2,\cdots,n;\ n\in Z^{+})$，“阶段入侵检测博弈”将被重复地进行。为简化起见，本章假设“成员传感器节点”S 与“簇头入侵检测代理”R 在“阶段博弈”t_k 和“阶段博弈”t_{k-1} 具有相同的支付矩阵，也就是说，在“多阶段动态入侵检测博弈”中不存在收益的折扣现象。

根据贝叶斯规则，“簇头入侵检测代理”R 可以从“阶段博弈”t_{k-1} 更新得到“阶段博弈”t_k 的“推断”值。设 $h_S(t_k)$ 为“成员传感器节点”S 的历史动作，$a_S(t_k)$ 为“成员传感器节点”S 在“阶段博弈”t_k 的动作，$p(\theta_S=1|a_S(t_k),\ h_S(t_k))$ 为“后验推断”。这里的“后验推断”表示在“阶段博弈”t_k“成员传感器节点”是恶意节点的概率。

定义 3-2 “簇头入侵检测代理”R 的“后验推断”的计算式为

$$p(\theta_S=1\mid a_S(t_k),h_S(t_k))=\frac{p(\theta_S=1\mid h_S(t_k))\cdot p(a_S(t_k)\mid \theta_S=1,h_S(t_k))}{\sum\limits_{\theta'_S\in\Theta_S}p(\theta'_S\mid h_S(t_k))\cdot p(a_S(t_k)\mid \theta'_S,h_S(t_k))} \quad (3\text{-}16)$$

式中，$p(\theta_S|h_S(t_k))$ 为在历史动作 $h_S(t_k)$ 下的“先验推断”；$p(a_S(t_k)|\theta_S,\ h_S(t_k))$ 为在“阶段博弈”t_k 中的“成员传感器节点”S 在采取历史动作 $h_S(t_k)$ 的前提下选择动作 $a_S(t_k)$ 的概率。

由于任何的“簇头入侵检测代理”R 都存在检测率和误报率，“簇头入侵检测代理”R 从观测到的“成员传感器节点”S 的动作中不一定能正确地反映实际的入侵检测现状。因此，本章在计算后验概率 $p(a_S(t_k)|\theta_S,\ h_S(t_k))$ 时，将考虑检测率和误报率的影响，这些将分别由以下的式子得到，即

$$p(\text{Attack}\mid\theta_S=1,\ h_S(t_k))=\alpha\rho+\beta(1-\rho), \quad (3\text{-}17)$$

$$p(\text{Cooperate}\mid\theta_S=1,\ h_S(t_k))=(1-\alpha)\rho+(1-\beta)(1-\rho), \quad (3\text{-}18)$$

$$p(\text{Attack}\mid\theta_S=0,\ h_S(t_k))=\beta, \quad (3\text{-}19)$$

$$p(\text{Cooperate}\mid\theta_S=0,h_S(t_k))=1-\beta, \quad (3\text{-}20)$$

式中，$1-\alpha$ 为负检测率；$1-\beta$ 为正误报率。

定义 3-3 “多阶段动态入侵检测博弈”是一个五元组$\mathbb{M}=(N,\Theta,A,P(D),U)$，其中：

- N、Θ、A 和 U 的定义与定义 3-1 中的 N、Θ、A 和 U 相同。
- $P(D)=(p(\theta_S=1|h_S(t_k)),\ 1-p(\theta_S=1|h_S(t_k)))$，其中 $p(\theta_S=1|h_S(t_k))$ 表示“成员传感器节点”在阶段 t_k 采取历史动作 $h_S(t_k)$ 的前提下是恶意节点的概率，在“阶段博弈”t_k 结束时，其值将根据式(3-16)计算得到的 $p(\theta_S=1|a_S(t_k),\ h_S(t_k))$ 进行更新。

随着“推断”值的更新，“多阶段动态入侵检测博弈”将以序贯的方式进行，最后通过“完美贝叶斯均衡”表示“多阶段动态入侵检测博弈”的均衡。在整个博弈过程中，“成员传感器节点”S 和“簇头入侵检测代理”R 为最大化它们各自的效用，并不总是在每一个阶段博弈中采用相同的策略，并且随着“多阶段入侵检测动态博弈”的进行，它们的最优响应策略与可能

改变的当前“推断”值相互独立。接下来，本章讨论如何得到“成员传感器节点”S 的类型的“推断”值，并将利用得到的完美贝叶斯均衡得到“成员传感器节点”S 和“簇头入侵检测代理”R 的最优响应策略。在寻找“多阶段动态入侵检测博弈”的完美贝叶斯均衡之前，本章首先要说明该博弈模型满足必需的贝叶斯条件。

定义 3-4[15]　贝叶斯条件包括：

B(i)　“后验推断”是相互独立的，并且参与者 i 的所有类型具有相同的“先验推断”。

B(ii)　“先验推断”到“后验推断”的更新通过贝叶斯规则实现。

B(iii)　参与者不传递任何参与者所不知道的事情信号。

B(iv)　“后验概率”在 Θ 上的共同的联合概率分布是一致的。

引理 3-1　“多阶段动态入侵检测博弈”满足贝叶斯条件。

证明　因为“簇头入侵检测代理”R 只有一种类型，因此 B(i)满足。因为式(3-16)由贝叶斯规则得到，因此 B(ii)满足。因为“成员传感器节点”S 的信号由它的动作来决定，并且如果条件 $a_S(t_k)=a'_S(t_k)$ 成立，那么 $p(\theta_S|a_S(t_k),\ h_S(t_k))=p(\theta_S|a'_S(t_k),\ h_S(t_k))$，因此 B(iii)满足。由于“多阶段动态入侵检测博弈”在任何阶段只有两个参与者，并且没有其他的参与者会影响“簇头入侵检测代理”R 对“成员传感器节点”S 的“推断”值的更新，因此 B(iv)满足。证毕。

定理 3-3　“多阶段动态入侵检测博弈”存在混合策略完美贝叶斯均衡。

证明　在“阶段博弈”t_k，设“恶意成员传感器节点”在“阶段博弈”t_k的策略为

$$\sigma_{S_k}=(\rho_k,1-\rho_k) \tag{3-21}$$

式中，ρ_k为“恶意成员传感器节点”$\theta_S=1$ 选择动作 Attack 的概率。设“簇头入侵检测代理”R 在“阶段博弈”t_k的策略为

$$\sigma_{R_k}=(\delta_k,1-\delta_k) \tag{3-22}$$

式中，δ_k为“簇头入侵检测代理”R 采取动作 Defend 的概率。对“簇头入侵检测代理”R 而言，在“阶段博弈”t_k采取动作 Defend 和 Idle 的期望收益分别是

$$\begin{aligned}Eu_R(\text{Defend})=&\rho_k p(\theta_S=1\mid h_S(t_k))(\alpha g_D-(1-\alpha)g_A-c_D)\\&+(1-\rho_k)p(\theta_S=1\mid h_S(t_k))(-\beta l_F-c_D)\\&+(1-p(\theta_S=1\mid h_S(t_k)))(-\beta l_F-c_D)\end{aligned} \tag{3-23}$$

和

$$\begin{aligned}Eu_R(\text{Idle})=&\rho_k p(\theta_S=1\mid h_S(t_k))(-g_A)+(1-\rho_k)p(\theta_S=1\mid h_S(t_k))\cdot 0\\&+(1-p(\theta_S=1\mid h_S(t_k)))\cdot 0\\=&-\rho_k p(\theta_S=1\mid h_S(t_k))g_A\end{aligned} \tag{3-24}$$

在“恶意成员传感器节点”$\theta_S=1$ 采取最优混合策略 $\sigma^*_{S_k}$ 的情况下，由“簇头入侵检测代理”R 采取动作 Defend 和 Idle 的无差异性可以得到

$$Eu_R(\text{Defend})=Eu_R(\text{Idle}) \tag{3-25}$$

因此，“簇头入侵检测代理”R 的最优混合策略为

$$\rho_k^*=\frac{\beta l_F+c_D}{p(\theta_S=1\mid h_S(t_k))(\alpha g_D+\alpha g_A+\beta l_F)} \tag{3-26}$$

对“恶意成员传感器节点”$\theta_S=1$ 而言，采取动作 Attack 和 Cooperate 的期望收益分别是

$$\begin{aligned}Eu_S(\text{Attack})=&\delta_k p(\theta_S=1\mid h_S(t_k))((1-\alpha)g_A-\alpha g_D-c_A)\\&+(1-\delta_k)p(\theta_S=1\mid h_S(t_k))(g_A-c_A)\end{aligned} \tag{3-27}$$

和

$$
\begin{aligned}
Eu_S(\text{Cooperate}) =& \delta_k p(\theta_S = 1 \mid h_S(t_k))(g_C - c_C) \\
&+ (1-\delta_k) p(\theta_S = 1 \mid h_S(t_k))(g_C - c_C) \\
&+ \delta_k (1 - p(\theta_S = 1 \mid h_S(t_k)))(g_C - c_C) \\
&+ (1-\delta_k)(1 - p(\theta_S = 1 \mid h_S(t_k)))(g_C - c_C)
\end{aligned} \tag{3-28}
$$

在“簇头入侵检测代理”R 采取最优混合策略 $\sigma_{R_k}^*$ 的情况下，由“恶意成员传感器节点”$\theta_S=1$ 采取动作 Attack 和 Cooperate 的无差异性可以得到

$$
Eu_S(\text{Attack}) = Eu_S(\text{Cooperate}) \tag{3-29}
$$

因此，“恶意成员传感器节点”$\theta_S=1$ 的最优混合策略为

$$
\delta_k^* = \frac{p(\theta_S = 1 \mid h_S(t_k)) g_A - p(\theta_S = 1 \mid h_S(t_k)) c_A - g_C + c_C}{p(\theta_S = 1 \mid h_S(t_k))(\alpha g_A + \alpha g_D)} \tag{3-30}
$$

综上所述，在“阶段博弈”t_k 存在混合策略完美贝叶斯均衡$(\sigma_{S_k}^*, \sigma_{R_k}^*)$，其中 $\sigma_{S_k}^*$ 和 $\sigma_{R_k}^*$ 分别与检测率 α、误报率 β 和“后验推断”$p(\theta_S=1|h_S(t_k))$有关。证毕。

定理 3-3 表示在“多阶段动态入侵检测博弈”中，两个理性参与者“成员传感器节点”S 和“簇头入侵检测代理”R 将选择最优策略对 $(\sigma_{S_k}^*, \sigma_{R_k}^*)$。随着“多阶段动态入侵检测博弈”的进行，它们将各自根据定理 3-3 选择最优的动作以获取最大的利益。

3.3.5 基于完美贝叶斯均衡的入侵检测机制设计

根据“多阶段动态入侵检测博弈”的完美贝叶斯均衡，本章提出并设计了一种适合于无线传感器网络的入侵检测机制。图 3-3 给出了“成员传感器节点”S 和“簇头入侵检测代理”R 在入侵检测过程中进行的动作交互。

在图 3-3 中，基于完美贝叶斯均衡的入侵检测机制包括 4 个部分：存储数据区、管理者、“成员传感器节点”S 和“簇头入侵检测代理”R。存储数据区主要用于存储“多阶段动态入侵检测博弈”涉及的参数 g_D、g_A、c_A、c_C、c_D、l_F、α、β 和 $p(\theta_S=1|h_S(t_k))$。因为“成员传感器节点”S 可能是恶意或正常的节点，所以它可能会采取动作 Attack 或 Cooperate，这些动作所产生的信息将会形成监控数据并发送到“簇头入侵检测代理”R。在“簇头入侵检测代理”R 开始工作之前，管理员首先已配置好“簇头入侵检测代理”R 的相应参数以便尽可能地使它工作得更加可靠和准确。在“簇头入侵检测代理”R 中，入侵检测引擎能利用已有的异常和误用检测技术判断监控数据是恶意的还是正常的。然后“簇头入侵检测代理”R 从存储数据区获得相应的博弈参数并初始化博弈模型，从而建立“阶段入侵检测博弈”。该博弈模型将接收来自入侵检测引擎中的输出数据和由管理者根据经验值设定的支付矩阵。其中，计算“簇头入侵检测代理”R 要采取动作 Defend 的概率 δ_k^* 需要来自“阶段入侵检测博弈”的数据支持，其计算过程对整个入侵检测机制而言是一个关键步骤，这是因为根据定理 3-3，得到 δ_k^* 就能确定“簇头入侵检测代理”R 将以何概率选择动作 Defend 和 Idle。最后，“簇头入侵检测代理”R 将计算 $p(\theta_S=1|a_S(t_k), h_S(t_k))$，并据此更新 $p(\theta_S=1|h_S(t_k))$后存入存储数据区，以备下一“阶段入侵检测博弈”使用。这样经过反复迭代，形成的“多阶段动态入侵检测博弈”被用于决定“簇头入侵检测代理”R 何时启动的策略。

下面给出基于完美贝叶斯均衡的无线传感器网络入侵检测代理何时启动最优策略算法。

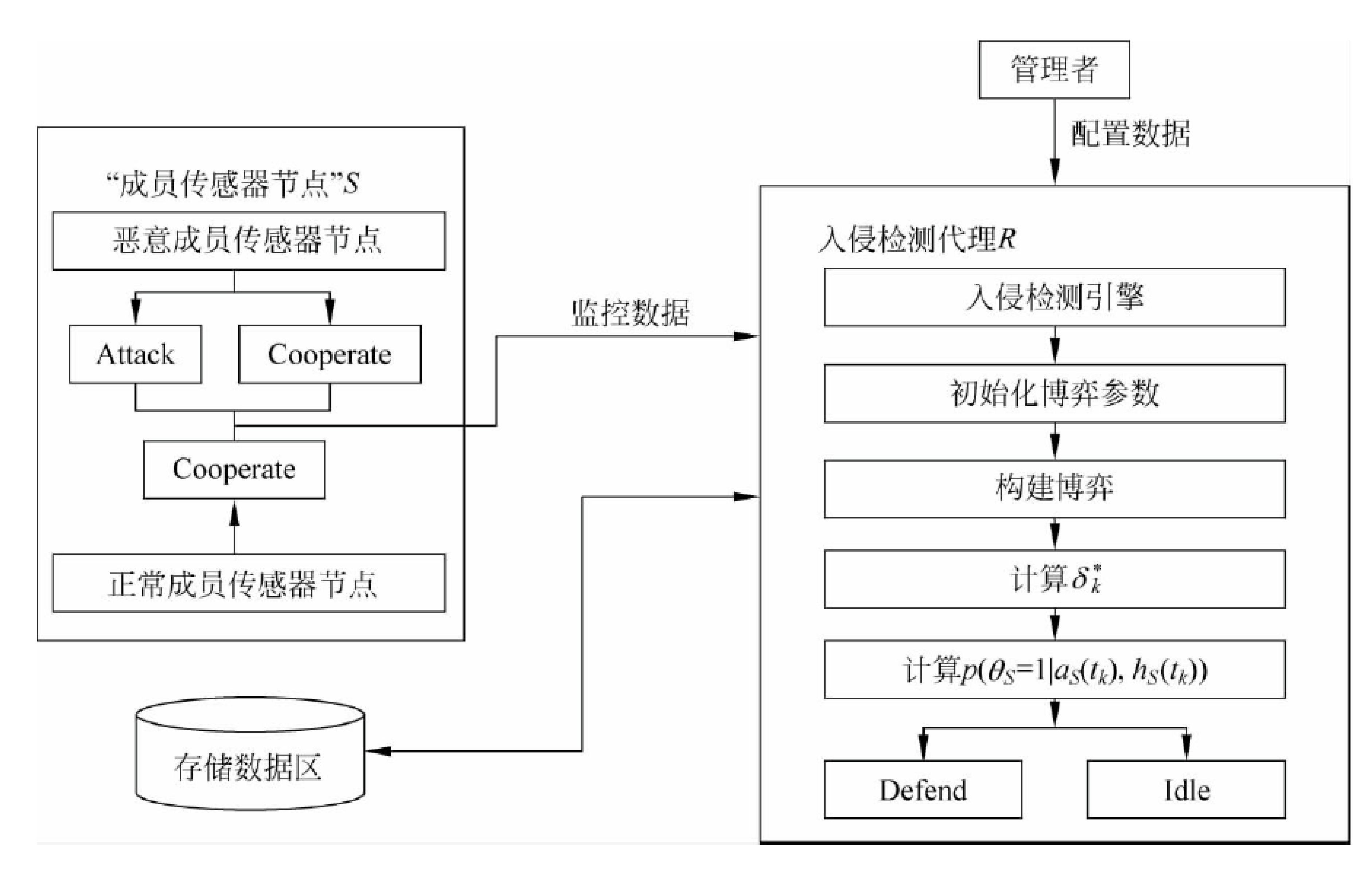


图 3-3　基于完美贝叶斯均衡的入侵检测机制

算法 3-1　无线传感器网络入侵检测代理何时启动最优策略算法。

```
"簇头入侵检测代理"R 选择动作 Idle。
当有监控数据到达时，唤醒"簇头入侵检测代理"R。
Do WHILE .T.
  "簇头入侵检测代理"R 从监控数据中获取一条记录。
  IF 记录包含恶意信息 THEN
    IF "阶段入侵检测博弈"不存在 THEN
      根据给定的博弈参数，建立"阶段入侵检测博弈"。
    ELSE
      获取"阶段入侵检测博弈"。
    ENDIF
    根据式(3-30)计算 $\delta_k^*$ 。
    根据式(3-16)计算 $p(\theta_S=1| a_S(t_k), h_S(t_k))$。
    由 $p(\theta_S=1| a_S(t_k), h_S(t_k))$更新 $p(\theta_S=1|h_S(t_k))$。
    将 $p(\theta_S=1|h_S(t_k))$ 存储到存储数据区。
    "簇头入侵检测代理"R 以概率 $\delta_k^*$ 选择动作 Defend。
  ELSE
    "簇头入侵检测代理"R 选择 Idle。
  ENDIF
  IF 监控数据处理结束
    EXIT;
  ENDIF
ENDDO
```

3.4 实验

本章利用 MATLAB 2010a 描述“多阶段动态入侵检测博弈”的参数并实现相应的模拟实验。在“阶段博弈”t_k，当检测率α和误报率β分别选择不同的值时，将比较用于确定“恶意成员传感器节点”类型的“后验推断”$p(\theta_S=1|\ a_S(t_k),\ h_S(t_k))$的变化情况。另外，本章还根据$\alpha$、$\beta$和$p(\theta_S=1|h_S(t_k))$的变化情况揭示$\rho_k^*$和$\delta_k^*$的变化趋势，它们将分别决定“恶意成员传感器节点”选择动作 Attack 的概率和决定“簇头入侵检测代理”R选择动作 Defend 的概率。为实现这些实验，根据无线传感器网络通常状况下的经验值，本章假设相应的博弈参数值如下：$g_A=250$、$g_C=5$、$g_D=200$、$c_A=20$、$c_C=5$、$c_D=10$ 和 $l_F=15$。

图 3-4 给出了在相同的误报率$\beta=0.05$前提下，当检测率α变化时，“簇头入侵检测代理”R计算得到的“后验推断”的收敛速度。从中可以看出，检测率α值越大，“后验推断”值收敛到 1 的速度越快。例如，当$\alpha=0.9$时，需要 10 次“阶段博弈”，“后验推断”$p(\theta_S=1|\text{Attack},\ h_S(t_k))$值收敛到 1；当$\alpha=0.7$时，需要 12 次“阶段博弈”，“后验推断”$p(\theta_S=1|\text{Attack},\ h_S(t_k))$值收敛到 1；当$\alpha=0.5$时，就需要 16 次“阶段博弈”才使“后验推断”$p(\theta_S=1|\text{Attack},\ h_S(t_k))$值收敛到 1。

图 3-5 考虑在相同的检测率$\alpha=0.9$前提下，不同的误报率对“簇头入侵检测代理”R计算得到的“后验推断”的影响。从中可以看出，误报率β值越小，“后验推断”值收敛到 1 的速度就越快。例如，当$\beta=0.01$时，需要 5 次“阶段博弈”使“后验推断”值收敛到 1；当$\beta=0.02$时，需要 7 次“阶段博弈”使“后验推断”值收敛到 1；而当$\beta=0.05$时，就需要 10 次“阶段博弈”才能使得“后验推断”值收敛到 1。根据图 3-4 和图 3-5 的实验结果，检测率α值的变大和误报率β值的变小都将使“簇头入侵检测代理”R判断“成员传感器节点”S是否为恶意节点的速度变快。也就是说，“簇头入侵检测代理”R检测“恶意成员传感器节点”的收敛速度将随着检测精度的提高而变快。

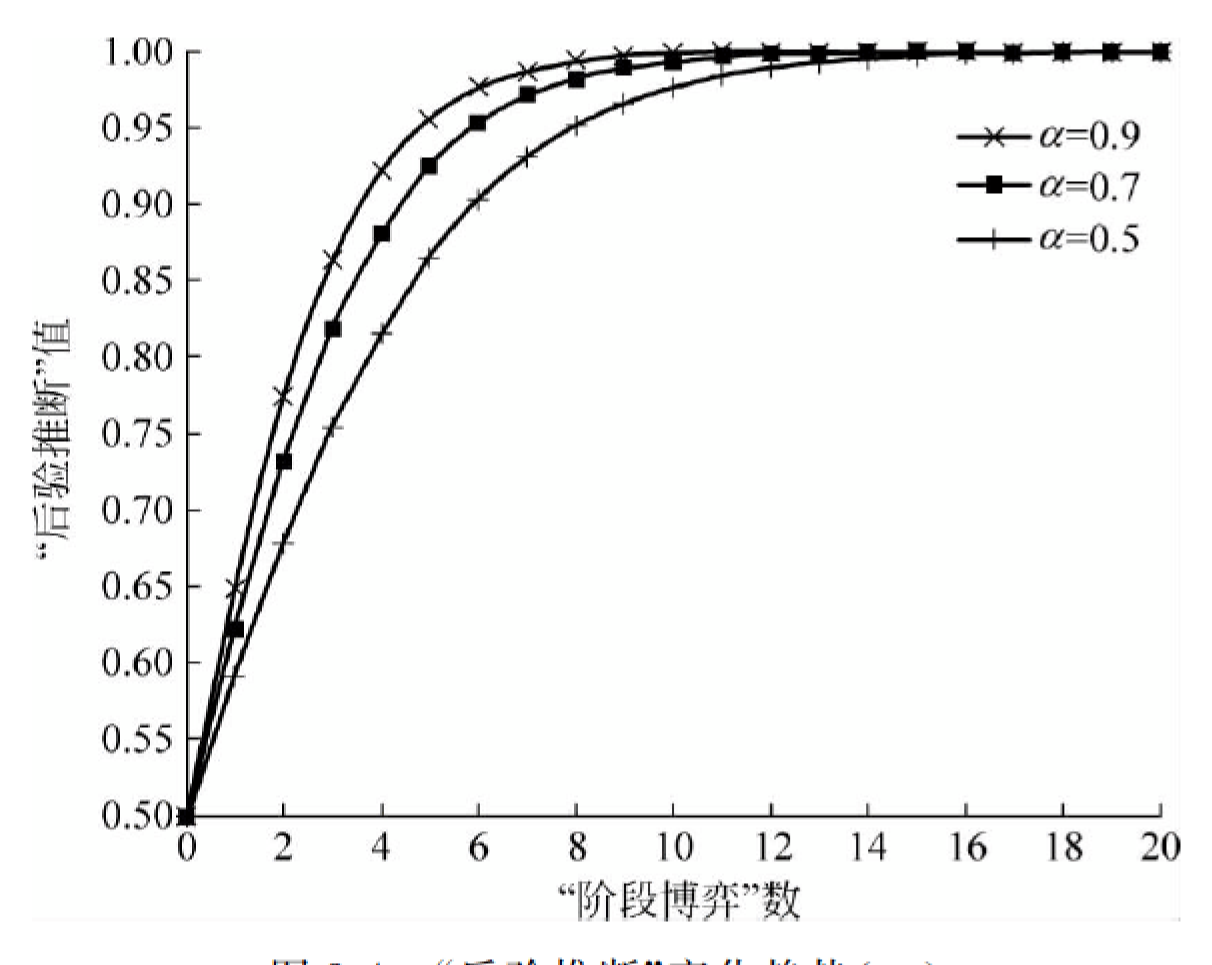


图 3-4 “后验推断”变化趋势(一)

为了观察“多阶段动态入侵检测博弈”的完美贝叶斯均衡变化趋势，假设“簇头入侵检测代理”R具有固定的误报率$\beta=0.05$，初始状态下“簇头入侵检测代理”R的检测率$\alpha=0.6$，

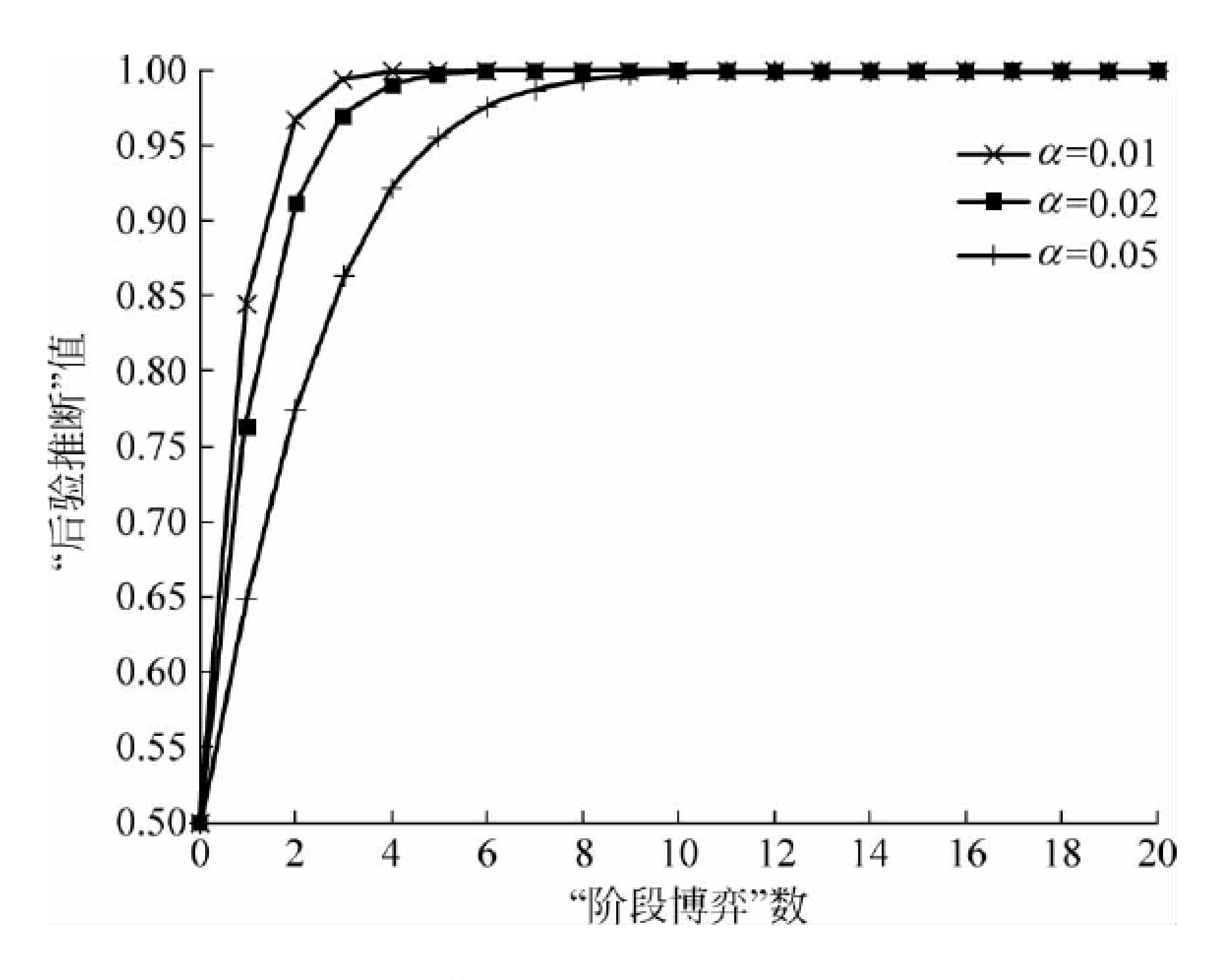


图 3-5　“后验推断”变化趋势(二)

“恶意成员传感器节点”选择动作 Attack 的概率 $\rho_k=0.5$,“簇头入侵检测代理”R 选择 Defend 的概率 $\delta_k=0.9$。由式(3-26)和式(3-30),“多阶段动态入侵检测博弈”的“完美贝叶斯均衡对”($\sigma_{S_k}^*,\sigma_{R_k}^*$)除与上述假定的博弈参数、“恶意成员传感器节点”和“簇头入侵检测代理”的期望收益有关外,还跟“簇头入侵检测代理”R 计算得到的“后验推断”有关。图 3-6 给出了当检测率 α 从 0.6 变化到 1 时,“恶意成员传感器节点”选择 Attack 和“簇头入侵检测代理”R 选作动作 Defend 的概率变化趋势。从中可以看出,α 值越大,ρ_k^* 和 δ_k^* 的值越小,这意味着“恶意成员传感器节点”选择动作 Attack 和“簇头入侵检测代理”R 选择动作 Defend 的概率随着 α 值的变大而变小。因此,“簇头入侵检测代理”R 应该努力提高它的检测率以降低用于防御“恶意成员传感器节点”攻击行为的成本。

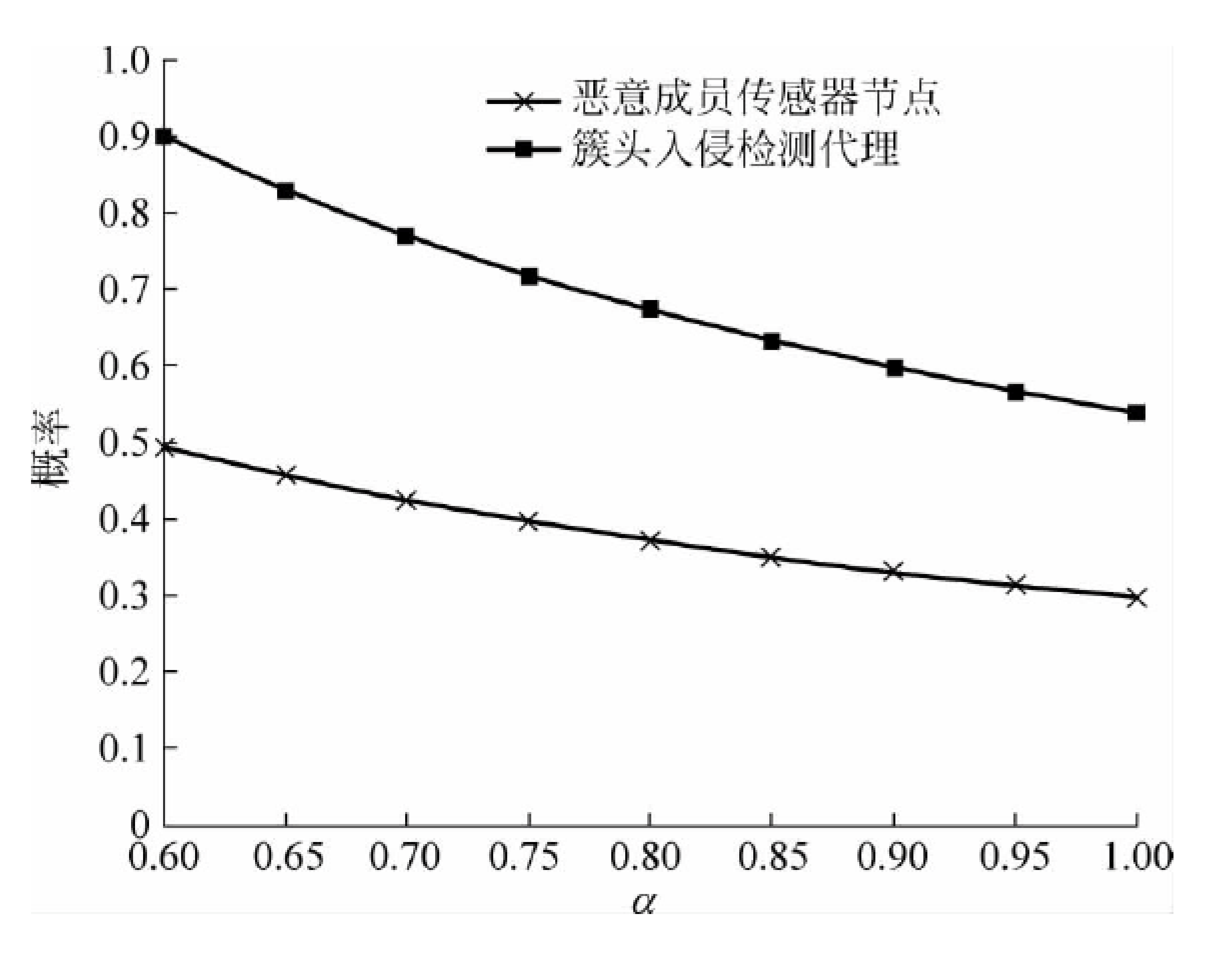


图 3-6　概率变化趋势

从上述实验结果可知,“多阶段动态入侵检测博弈”能为“簇头入侵检测代理”提供优化的动作选择策略,能根据“恶意成员传感器节点”的历史动作合理地更新用于推测“恶意成员传感器节点”的“后验推断”值。通过提高“簇头入侵检测代理”的检测率和降低它的误报率,

“簇头入侵检测代理”判断“成员传感器节点”是否为恶意节点的收敛速度将明显变快，从而，使得基于完美贝叶斯均衡的入侵检测机制能更快且更主动地防御“恶意成员传感器节点”的攻击行为。

3.5 小结

为了选择最优的入侵检测响应策略以节省无线传感器网络入侵检测系统的资源消耗，本章提出了一种基于信号博弈的入侵检测博弈模型。该博弈模型能反映“成员传感器节点”和“簇头入侵检测代理”之间的交互，并能揭示“成员传感器节点”和“簇头入侵检测代理”在不同的“阶段博弈”中如何采取攻击策略和防御策略的规律。本章选择的分布—集中混合式网络结构能有效地降低“簇头入侵检测代理”执行时的能量消耗。在每个独立的阶段，各“阶段入侵检测博弈”能很好地展示传感器节点之间的交互行为，得到的纯策略和混合策略贝叶斯均衡使“簇头入侵检测代理”知道何时选择动作 Defend 或 Idle，也就是说，“簇头入侵检测代理”不必总是选择 Defend，从而有效地降低用于运行“簇头入侵检测代理”的能量消耗。随着博弈的进行，本章将“阶段入侵检测博弈”转变成能根据“恶意成员传感器节点”的当前和历史行为进行动态更新“推断”值的“多阶段动态入侵检测博弈”模型，并且得到了能使“簇头入侵检测代理”使用最佳响应策略的混合策略完美贝叶斯均衡。根据这些完美贝叶斯均衡提出的入侵检测算法实现了入侵检测博弈的应用。实验结果说明了提出的入侵检测博弈模型在预测“成员传感器节点”类型方面的有效性，从而“簇头入侵检测代理”能主动地选择优化的策略防御“恶意成员传感器节点”的攻击。

第 4 章

基于演化博弈的无线传感器网络节点信任演化动力学研究

本章利用演化博弈研究传感器节点间的信任决策过程，根据各个传感器节点能选择不同策略的实际情况建立“无线传感器网络信任博弈”模型，通过整合激励机制参数来说明激励机制对传感器节点选择动作的影响，使用复制动态动力学方程探索博弈模型的演化稳定策略，从而揭示无线传感器网络各传感器节点间的信任演化原理。

4.1 引言

当前，无线传感器网络的应用范围大致可概括为监测（Monitoring）和追踪（Tracking）两大领域[5]。监测领域主要包括环境监测、健康监测、电网监测、工业自动化监测、地震监测等，追踪领域主要涉及目标跟踪，如动物跟踪、特定人员跟踪、车辆跟踪等。为了支持这些应用，如何保证无线传感器网络的安全是首先要解决的问题。采用基于密码学的方法是一种传统的网络安全技术，这种技术常称为硬安全（Hard Security）[42]，主要解决数据机密性、数据完整性和传感器节点的身份认证等问题。另一种是采用称为软安全（Soft Security）[42]的信任和声誉管理系统，这种机制能有效地应对部分已通过硬安全检查的正常节点为了自己能获取更大的利益而提供错误或虚假信息的欺骗行为。因此，信任和声誉管理系统是硬安全技术的有效补充，对保障无线传感器网络正常安全地运行具有重要作用。

作为最近几年应用到无线传感器网络中的安全技术，基于信任的安全机制有其自身的特点。与加解密、监测非法入侵等安全技术相比，信任机制是在以无线传感器网络传感器节点为中心的环境中，帮助各传感器节点建立信心，推动传感器节点之间的协作，降低与其他传感器节点合作的风险。通常，一个无线传感器网络信任和声誉管理系统需要解决如何确定传感器节点的信任度及传感器节点之间的信任如何进行演化的问题。这是由于一旦在无线传感器网络中部署信任管理系统后，传感器节点间的信任决策及其动力学演化将决定传感器节点间是否采取合作的行为，从而影响整个无线传感器网络的稳定和安全。在这个过程中，无线传感器网络信任管理系统需要收集并存储用于确定信任度的相应凭据，需要记录传感器节点的动作行为，再计算得到对应传感器节点的信任度值。根据其他传感器节点的信任度值，一个传感器节点将通过信任决策确定是否与其他传感器节点进行合作，使得那些有欺骗行为（相应的信任度要低）的传感器节点在试图进行通信时得不到其他节点的合

作通信支持,从而减少传感器节点的欺骗行为,保障无线传感器网络的正常通信。因此,研究无线传感器网络传感器节点的信任决策和信任演化动力学过程对保障无线传感器网络的稳定和安全起重要作用[111]。

博弈论作为研究参与者之间理性决策的有效工具[112],已广泛用于无线传感器网络优化的多个方面,如能量消耗优化[113]、功率优化分配[114]等问题。使用博弈论工具,能得到参与者采取策略的均衡点,从而决定各参与者的最优行动。与传统博弈类型不同的是,演化博弈认为一个种群(Population)中的参与者具有进化能力并能为满足自身利益而重复地进行博弈,直至整个种群达到一定程度的均衡。根据演化博弈理论,一个种群的动力学演化反映的是那些选择具有较高收益策略的个体(Individual),将逐步增加它们在整个种群中的比例,而那些选择具有较低收益策略的个体比例将逐步减少。通过使用演化博弈这种工具,可以分析一个种群中各参与者对不同策略进行的选择演化,从而揭示整个种群在假定利益前提下的决策动力学演化规律。

本章将利用演化博弈研究传感器节点间的信任决策过程,从而揭示无线传感器网络各传感器节点间的信任演化原理。根据演化博弈论的特点,本章将整个无线传感器网络看作一个种群,并将每个传感器节点看作这个种群中的一个个体。然后根据各个传感器节点能选择不同策略的实际情况建立"无线传感器网络信任博弈"模型,并且为了研究激励机制对传感器节点选择动作 Trust(即可以合作通信)的影响,在"无线传感器网络信任博弈"模型中整合激励机制参数。为了说明"无线传感器网络信任博弈"模型的稳定性,通过复制动态动力学方程探索"无线传感器网络信任博弈"的演化稳定策略。

在扩展作者前期工作[115]的基础上,本章的工作主要包括以下内容:

(1) 建立适用于传感器节点信任决策的"无线传感器网络信任博弈"模型,该模型能在传感器节点进行信任决策时正确地反映传感器节点选择不同动作的收益情况。

(2) 在"无线传感器网络信任博弈"模型中绑定激励机制参数后,也就是说,若在实际的无线传感器网络信任管理系统中引入激励机制,能有效地减少传感器节点选择动作 Distrust(即对相互通信采取不合作的行为)的比例,从而使各传感器节点向选择动作 Trust 转化,达到改善无线传感器网络的稳定性和安全性的目的。

(3) 得到与"无线传感器网络信任博弈"相关的演化稳定策略定理,这些定理给出了达到演化稳定策略的条件,并且能为无线传感器网络信任管理系统的实际设计提供理论基础。

本章其余章节安排如下:4.2 节说明相关工作;4.3 节建立"无线传感器网络信任博弈"模型并利用复制动态方程寻求相应的演化稳定策略及揭示传感器节点信任演化机理;4.4 节通过实验验证"无线传感器网络信任博弈"的演化稳定策略和激励机制的有效性;4.5 节给出本章小结。

本章涉及的符号含义如下:

s_i表示第 i 个纯策略。

S 表示给定种群中各个体可选择的纯策略集合。

$\phi_i(t)$表示在时刻 t 选择纯策略 s_i的个体数量。

$\theta_i(t)$表示在时刻 t 选择纯策略 s_i的个体比例。

$\theta(t)$表示种群在时刻 t 的混合策略。

$u(s_i,\theta(t))$表示个体在时刻 t 选择纯策略 s_i的期望收益。

$\bar{u}(\theta(t),\theta(t))$表示种群的平均期望收益。

T 表示一个传感器节点的信任度。

α 表示一个传感器节点选择某个动作后产生的收益和它的信任度之间的调节因子。

G_T表示一个传感器节点选择动作 Trust 后得到的收益。

G_D表示一个传感器节点选择动作 Distrust 后得到的收益。

G_C表示一个传感器节点在相邻节点选择 Trust 时所得到的合作收益。

C 表示一个传感器节点在发送自身数据或转发其他传感器节点的数据时所产生的成本。

L 表示一个传感器节点在相邻传感器节点选择动作 Distrust 时对其造成的损失。

$\mathbb{G}=(P,N,S,U)$表示“无线传感器网络信任博弈”。

θ 表示选择动作 Trust 的传感器节点在整个无线传感器网络中所占的比例。

4.2　相关工作

信任就是相信对方，是一种建立在自身知识和经验基础上的判断，是一种实体与实体之间的主观行为，是基于观察所得到的经验总结。信任能识别无线传感器网络中的恶意节点和自私节点，被认为是对基于密码体制安全措施的有效补充。

虽然信任机制近几年来才被引入到无线传感器网络，但研究人员对无线传感器网络信任机制研究的关注度很高，已有大量文献发表。荆琦等人[116]综述了无线传感器网络环境下信任管理的特点、分类方法、框架设计等，并介绍了无线传感器网络下的典型信任管理系统。指出信任管理系统的核心是以信任计算模型为中心的信任管理框架设计，讨论了信任要素、信任计算模型和信任值的应用 3 个方面。Momani[117]综述了无线传感器网络中已提出的主要信任模型。Yu 等人[111]和 Esch[118]分别综述了包括无线传感器网络在内的无线通信领域中的信任和信誉管理系统。Lopez 等人[119]给出了在无线传感器网络实施信任管理系统的最佳实践。Yu 等人[120]分析了无线传感器网络中与信任模式相关的攻击类型与对策，从路由安全和数据安全两方面总结了无线传感器网络环境下信任管理的研究现状，提出了进一步的研究发展方向。

目前，对无线传感器网络信任的研究主要集中在对节点进行信任度计算方面，再将信任度值应用于路由[121-123]、数据融合[124]、安全架构[125-130]等无线传感器网络的基础支撑技术中，从而全面提高无线传感器网络的安全性和可用性。莫英红等人[124]将传感器节点的信任分为传感信任、传递信任和融合信任，利用局部相关一致性原理检测传感器节点的功能行为，提出一种按功能行为进行分类信任的安全数据融合方法。这种方法可以有效地提高无线传感器网络数据融合的安全性和可靠性，从而延长网络寿命。黄廷磊和李小龙[127]针对无线传感器网络中正常节点误判为恶意节点的问题，提出一种基于本地信息评估传感器节点信任值的信任管理机制。王建萍等人[125]把声誉和组信任结合用于无线传感器网络中的实体认证协议，通过引入对称密钥体制，避免了非对称密钥体制实施数字签名计算量大的问题，有效地降低了协议的认证时延。吕林涛等人[121]提出了一种面向分层路由的信任模型用于解决无线传感器网络不能有效地检测出内部恶意节点攻击所引发的网络安全问题，该模型能发现并排除来自无线传感器网络内部实施攻击的恶意节点，从而提高安全性能。董

慧慧和郭亚军[128]将传感器节点通信、数据和能量相结合，把感知数据和传感器节点能量加入到传感器节点的信任评估中，来计算各自的信任值，这种基于多角度的信任模型能够更准确、简单地判断一个传感器节点的可信性，从而建立一个传感器节点之间相互可信的无线传感器网络。张乐君等人[131]建立了基于社会网络关联度的无线传感器网络节点信任模型，提出了基于关联度的传感器节点信誉度的计算方法，并设计了基于滑动窗口的传感器节点信任值计算及更新算法。Maarouf 等人[122]提出了一种基于声誉系统的信任感知路由方案，利用概率论计算方法在保证满足恶意节点检测度的前提下降低了邻居节点的监测频率。Aivaloglou 和 Gritzalis[132]基于证书和行为的混合方法来建立传感器节点的信任度，使得传感器节点的信任度可以根据网络的配置改变而演化。Leligou 等人[123]将信任机制用于解决位置感知路由协议中的攻击问题，传感器节点将在考虑位置和相邻传感器节点信任值的基础上决定合适的路由。Zhan 等人[133]提出一种能对多种错误和攻击容忍的无线传感器网络信任管理系统，能根据历史数据和当前行为的风险利用高斯模型细粒度地计算出传感器节点相应的信任度。Boukerch 等人[126]为无线传感器网络提出了一种基于智能体(Agent)的信任和声誉管理框架，该框架具有较小的额外信息和时延。Mármol 和 Pérez[134]将蚁群系统中的信任建立机制应用于无线传感器网络，提出的基于生物启发技术的信任管理系统具有精确、鲁棒的特点。He 等人[129]针对无线医疗传感器网络(Wireless Medical Sensor Networks)中的安全和隐私需求，提出一种攻击容忍的轻量级信任管理系统，实现了对传感器恶意节点的有效检测。Bao 等人[130]通过来自整个网络的多维信任信息来评价一个传感器节点的信任度，设计了一个层次化的信任管理系统，并用于基于位置感知的路由协议和基于信任的入侵检测领域。Jiang 等人[135]针对当前信任度计算仅考虑通信行为的现状，提出一种新的直接信任度和推荐信任度计算方法，其中，直接信任度的计算考虑了节点通信、能耗、数据传输等因素，而推荐信任度的计算考虑了信任的可靠度和亲密度。Ren 等人[136]针对无人照料的无线传感器网络，提出一种有效且鲁棒的信任度计算和存储方法，其中，地理位置哈希表(Geographic Hash Table)被用于标识需要存储信任度的节点，显著降低了存储成本。

Chae 等人[137]针对现有信任管理中清偿模式(Redemption Scheme)不能区分暂时错误(Temporary Errors)和假装恶意行为(Disguised Malicious Behaviors)的问题，提出了一种新的信任管理模式，很好地解决了该问题。Zhou 等人[138]根据节点位置和其他节点的信任值，通过优化调度信任管理系统中的看门狗(Watchdog)任务，得到了一种能量高效的信任管理系统。由于云计算具有的强大数据存储和处理能力，可以处理无线传感器网络中的巨量感知数据，Zhu 等人[139]针对这种传感云结构，提出一种新的信任度和信誉度计算方法以及相应的管理系统，能实现云服务提供者和传感网络提供者的认证，帮助用户正确选择云服务提供者，并使云服务提供者能选择合适的传感网络提供者的功能。

演化博弈与无线网络的结合是当前学术界的研究热点之一，在无线网络的不同方面都已有演化博弈的相关应用。张国鹏等人[140]利用演化博弈提出一种能有效激励 Ad Hoc 无线网络节点参与数据中继转发协作的问题，以节点中继所需的能量开销与数据分组数为均衡点建立单阶段博弈模型，验证节点的自私性，然后扩展单阶段博弈为基于策略可转换的演化博弈，并提出了协作激励策略。刘凤鸣和丁永生[141]运用演化博弈理论，对 P2P 网络节点信任度的动力学方程进行求解，并运用复制动态模型分析了节点之间信任关系的演化趋势，

揭示了节点间相互信任的演化动力学规律。而项兴彬等人[142]将演化博弈应用于P2P网络环境下文件共享时节点的信任建立。Niyato和Hossain[143]为了解决异构网络中不同服务区域中用户如何分享有限带宽的问题，利用演化博弈建立了相应的博弈模型用于反映这种用户之间进行的带宽竞争现状，研究了用户对网络选择的动力学过程。Tembine等人[144]以Aloha无线网络中多路控制(Multiple-access Control)和CDMA(Code Division Multiple Access)无线网络中功率控制为背景，将扩展后的演化博弈用于研究任意多个用户之间的非合作交互行为。Komathy和Narayanasamy[145]利用演化博弈提出一种动态且分布式的框架用于研究自组织Ad Hoc网络中自私节点的行为演化动力学问题，并与AODV路由协议结合说明了提出的博弈模型在促进自私节点转向合作时所起的作用。Anastasopoulos等人[146]利用演化博弈中的复制动态动力学方程提出了一种能进行自适应调节编码和调制的机制实现最大化TCP吞吐量的目的。Wang等人[147]根据演化博弈具有的自适应学习能力和群体中的个体能在环境条件改变的情况下选择最优响应策略的特性，将其用于解决合作的频谱检测问题，得到了后续用户的行为演化动力学规律和相应的演化均衡策略。类似地，Wang等人[148]也利用演化博弈讨论了无线网络中的合作与共谋(Collusion)的问题。Chen等人[149]利用演化博弈提出一种适用于无线传感器网络环境的动态激励机制，促使传感器节点为最大化其适应度(Fitness)而动态调整策略，最终使自私节点能转向合作实现无线传感器网络的正常服务。另外，演化博弈的应用领域还包括处理无线传感器网络中的数据融合[150]、分析P2P网络中能实现节点间相互合作的自适应激励协议[151]、异构4G网络中为达到最优策略而进行的学习机制[152]等。

通过信任机制实现自治网络的安全过程可以看作是一个策略交互的过程，一个节点的信任决策将影响其他节点的信任决策(信任或不信任)和网络状态(安全或不安全)，这种决策的过程自然可以利用博弈论进行解释和分析。因此，在信任和博弈论的结合方面，已有一些不同的博弈类型被应用于不同网络环境下的信任机制研究。孟宪福和王动[153]通过引入惩戒机制，建立了一种基于重复博弈的P2P网络信誉模型，以达到激励节点协作的目的。罗俊海和范明钰[154]基于非合作完全信息静态博弈研究Ad Hoc网络中的节点行为，提出了相应的信任模型，每个节点根据自身的信誉度来获得资源，鼓励节点共享资源和转发数据，惩罚自私节点。黄宇等人[155]将完全信息扩展博弈运用于信任协商决策中，将自动信任协商过程转化为完全信息扩展博弈过程，构造了信任博弈树和支付函数，依据子博弈精炼纳什均衡来决定自动信任协商策略。刘继超等人[156]基于非合作完全信息静态博弈研究信任建立过程，根据纳什均衡解精减初始证书交换集，使得节点双方在获取对方最大信任的同时，自身隐私损失降到最低限度。陈晶等人[157]将整个信任系统分为证据收集、信任度量和服务博弈三部分，其中服务博弈模块利用非合作完全信息静态博弈分析度量结果，结合网络环境得到服务提供者行动的混合纳什均衡策略。孙玉星等人[158]通过重复概率博弈模型分析节点之间的信任推荐交互过程，给出了TFT(Tit For Tat)、GTFT(Generous TFT)、GT(Grim Trigger)、OT(One-step Trigger)等激励策略对提升节点间信任推荐协作的影响。桂劲松和吴敏[159]将信任、服务预测、非合作完全信息静态博弈相结合并应用于无线接入服务中，根据移动节点的历史行为信任等级和请求服务所需的资源来计算各个信任和请求服务等级的概率，再结合纳什均衡解给出无线接入点是否接纳的决策。Wang等人[160]将非合作完全信息静态博弈应用于P2P网络中资源请求者和资源提供者之间的信任协商过程，得

到的混合纳什均衡解被用于确定需要缓存的信任序列数量。Jaramillo 和 Srikant[161]针对 Ad Hoc 网络中的自私节点，利用演化博弈提出一种具有自适应特点的分布式声誉机制，有效避免了因错误地将正常节点误判为自私节点后该节点的报复行为。Mejia 等人[162]为了解决 Ad Hoc 网络节点间的合作问题，引入完全信息静态博弈建立节点间的信任模型，利用细菌算法使节点能快速学习到节点的合作行为，实现了在较短时间达到优化合作的目的。Yahyaoui[163]将竞价博弈和信任结合用于 Web 服务领域，在每次博弈中，不同的 Web 服务提供者对需要完成的任务以信任成本的方式进行报价，具有最低成本的 Web 服务提供者将赢得博弈。

但将博弈论方法应用于无线传感器网络信任研究的文献并不多见。杨东巍等人[164]为了帮助无线传感器网络做出既有利于自身收益又能抑制恶意节点的决策，提出了一种基于重复博弈能实现信任激励的时隙分配博弈模型。李紫川等人[165]针对无线传感器网络节点信任决策影响节点间互助转包问题，在考虑网络不可靠因素的基础上，引入节点反思机制，构建基于概率论方法的无线传感器网络节点信任演化模型，再通过动力学分析，推导出达到演化稳定状态的定理。Komathy 和 Narayanasamy[166]基于演化博弈构建了针对自私节点动态行为的模型并用于形成节点的信任值评估，但缺乏对节点信任值动态演化机理的深入分析。Agah 等人[44]利用合作博弈建立了关注节点合作度、信任度和安全质量三方面因素的模型，根据节点丢失的数据包率来判定节点的安全质量，并提出基于节点合作次数的信任度计算方法。Feng 等人[167]针对无线传感器网络信任管理系统忽视正常节点自私性和恶意节点非合作性的特点，基于贝叶斯博弈建立了未知类型节点和正常节点之间的博弈模型，给出了一种激励节点相互合作的机制。Duan 等人[168]基于博弈论提出了一种信任度计算框架，其中，给出的风险策略模型(Risk Strategy Model)有效促进了传感器节点之间的合作，博弈论在信任度获取过程中的应用降低了处理成本。Guo 等人[169]在利用离散粒子群优化方法建立节点并行联盟的基础上，为了最小化任务的执行时间和节点的能量消耗，基于博弈论设计了信任动态任务调度策略，提出了一种有效提高任务效率和网络可靠性的方法。

与上述相关工作相比，本章主要关注无线传感器网络中传感器节点间的信任决策过程，通过构建相应的无线传感器网络信任博弈模型来分析信任决策的演化动力学。其中无线传感器网络信任博弈模型考虑了传感器节点的信任度因素，能反映出传感器节点在信任决策过程中的利益得失。本章最后利用演化博弈中的复制动态动力学方程给出达到演化稳定策略的条件，这些结果将为无线传感器网络构建和设计信任管理系统提供理论基础。

4.3 无线传感器网络信任博弈

4.3.1 演化博弈与无线传感器网络信任的结合

无线传感器网络信任与各传感器节点的行为密切相关。无线传感器网络传感器节点一般处于不确定的环境中，具有相对变化的特点，各传感器节点通常根据各相邻节点之间的信任值进行相互之间的信任策略决策，这种由传感器节点选择信任策略到建立信任关系的过程将保证各节点之间的协作。各传感器节点通过与其他传感器节点的反复交互，不断地学习与模仿，来动态调整自身的信任或不信任策略，从而实现节点间的信任或不信任策略

选择的演化,最终达到信任或不信任策略的演化稳定,这个过程利用演化博弈中的复制动态动力学方程能很好地进行描述。因此,将演化博弈思想应用到无线传感器网络信任决策可以深刻地揭示信任的特征及演化机制,为提高无线传感器网络的安全性并促进无线传感器网络的稳定性提供理论基础。

无线传感器网络信任博弈的建立过程具有以下特点:

(1) 各传感器节点的行为具有有限理性。由于无线传感器网络信任博弈是多个传感器节点之间的博弈,即某个传感器节点在进行信任决策时,不仅要考虑其他传感器节点加入后对博弈的影响,还要考虑选择信任或不信任策略后自己和其他有利益关系的传感器节点之间的收益关系,而这些信息的处理体现了有限理性的特征。

(2) 无线传感器网络信任博弈具有重复性和非零和性。各传感器节点之间的博弈是重复进行的,双方的博弈过程都不会改变支付矩阵,彼此都可以看到对方过去的动作和收益。另外,当各传感器节点都选择信任策略时,双方的收益可以实现双赢,因此无线传感器网络信任博弈是非零和的。

(3) 无线传感器网络中各传感器节点进行策略决策时具有模仿性。当传感器节点不能完全正确地判断自己行为得失但知道前期利益相关参与者的收益得失的时候,模仿前期最佳动作就是它的最佳策略。本章后续内容将利用复制动态动力学方程分析这种策略选择的模仿性。

4.3.2 无线传感器网络信任博弈模型

无线传感器网络传感器节点信任建立过程中表现出的有限理性决定了个体不是一开始就能找到最优策略,它们会在博弈过程中不断学习,通过模仿与试错寻找较好的动作策略。同时,这种有限理性意味着无线传感器网络信任博弈的均衡是不断调整和改进的过程,而不是一次性选择的结果,而且即使达到了某个均衡也可能出现偏离的现象。实际上,对无线传感器网络信任博弈分析的核心不仅是个体的最优策略选择,还包括种群个体的策略调整过程、趋势和最终的稳定性,其中稳定性是指种群个体采用某个特定策略的比例不变,而非某个个体选择的策略不变。

定义 4-1 无线传感器网络信任博弈是一个由四元数组$G=(P,N,S,U)$组成的对称博弈,其中:

- P 表示由大量个体(传感器节点)组成的一个种群(无线传感器网络)。
- N 表示由传感器节点构成的个体集合。
- S 表示可供传感器节点选择的策略集合,其中 $S=\{s_1,s_2\}=\{\text{Trust},\text{Distrust}\}$。
- U 表示两个传感器节点在一次博弈中得到的收益形成的支付矩阵,其值如表 4-1 所示。

表 4-1　无线传感器网络信任博弈的支付矩阵

动作	Trust	Distrust
Trust	$G_T+G_C+\alpha T-2C$	$G_T+\alpha T-C-L$
Distrust	G_D+G_C-C	G_D

在表示无线传感器网络传感器节点的信任关系时,各传感器节点信任值常使用信任度

进行度量，文献[129，132，133，170]都给出了不同的信任度计算方法，本章不考虑如何计算传感器节点的信任度，但假设每个传感器节点都已具有某个信任度值，并且其信任度值越高表示越值得信任。

在无线传感器网络信任博弈中，每个传感器节点可以选择动作 Trust 或 Distrust。动作 Trust 意味着一个传感器节点和其他节点进行交互时与对方节点进行合作，而选择动作 Distrust 将导致与对方节点的合作失败。下面分别讨论各种情况。

情况 1 进行交互的两个传感器节点都选择动作 Trust。此时每个传感器节点都与对方传感器节点合作，帮助对方传感器节点转发数据包，从而提高了自身的信任度，所以每个传感器节点都得到了信任度收益 G_T。又因对方传感器节点选择动作 Trust 而帮助转发数据包得到了收益 G_C。另外，为激励传感器节点选择动作 Trust，选择动作 Trust 的传感器节点将得到 αT 的激励。同时，在发送自身或转发对方传感器节点数据包时导致能量消耗分别产生成本 C，因此，综合上述分析，每个传感器节点的总收益为 $G_T+G_C-2C+\alpha T$。

情况 2 一个传感器节点选择动作 Trust 而另一个传感器节点选择动作 Distrust。此时选择动作 Trust 的传感器节点因帮助对方传感器节点转发数据包得到信任度收益 G_T，并且得到 αT 的激励。同时，因转发对方传感器节点数据包产生成本 C，并且因对方传感器节点选择动作 Distrust，导致不合作使得自身数据包无法发送到目标传感器节点而产生损失 L。因此，选择动作 Trust 的传感器节点的总收益为 $G_T-C-L+\alpha T$。选择动作 Distrust 的传感器节点因为不需要为其他传感器节点转发数据包，所以节省了能量消耗和延长了生命期，从而获得了收益 G_D。同时因对方传感器节点选择动作 Trust 而帮助自己转发数据包得到了收益 G_C，但又因发送自身数据包产生成本 C。因此，选择动作 Distrust 的传感器节点的总收益为 G_D+G_C-C。

情况 3 两个传感器节点都选择动作 Distrust。此时将导致无线传感器网络的完全不合作，但因为不需要为其他传感器节点转发数据包，所以节省了能量消耗和延长了生命期，从而获得了收益 G_D。因此，两个传感器节点的总收益均为 G_D。

4.3.3 无线传感器网络信任演化稳定策略和动力学分析

由于无线传感器网络信任博弈模型中共包含 Trust 和 Distrust 两种动作，因此在由传感器节点构成的种群中，可设 $\theta(t)=(\theta_1(t),\theta_2(t))$ 表示种群在时刻 t 所处的混合策略，其中 $\theta_1(t)$ 表示选择动作 Trust 的传感器节点数比例，$\theta_2(t)$ 表示选择动作 Distrust 的传感器节点数比例，则有 $\theta_2(t)=1-\theta_1(t)$。为简化起见，以下记 $\theta_1(t)$ 为 θ。那么，由式(2-9)可得传感器节点在时刻 t 选择动作 Trust 的期望收益为

$$u(s_1,\theta(t))=\theta(G_T+G_C+\alpha T-2C)+(1-\theta)(G_T+\alpha T-C-L) \tag{4-1}$$

选择动作 Distrust 的期望收益为

$$u(s_2,\theta(t))=\theta(G_D+G_C-C)+(1-\theta)G_D \tag{4-2}$$

由式(2-10)得到整个无线传感器网络种群 P 的平均期望收益为

$$\bar{u}(\theta(t),\theta(t))=\theta u(s_1,\theta(t))+(1-\theta)u(s_2,\theta(t)) \tag{4-3}$$

因此，由式(2-12)可以得到传感器节点信任演化的复制动态动力学方程为

$$\begin{aligned}F(\theta)=\dot{\theta}&=\theta(u(s_1,\theta(t))-\bar{u}(\theta(t),\theta(t)))\\&=\theta(1-\theta)[\theta(G_T+\alpha T-G_D-C)+(1-\theta)(G_T+\alpha T-G_D-C-L)]\end{aligned} \tag{4-4}$$

令 $F(\theta)=0$，则复制动态动力学方程式(4-4)最多有 3 个稳定状态，即

$$\theta_1^* = 0 \tag{4-5}$$

$$\theta_2^* = 1 \tag{4-6}$$

$$\theta_3^* = (G_D + C + L - G_T - \alpha T)/L \tag{4-7}$$

其中式(4-7)表示的稳定状态可能与式(4-5)或式(4-6)表示的稳定状态相同。

根据演化稳定策略的性质，一个稳定状态必须对动态系统的微小扰动具有稳定性。这实际上和微分方程中的稳定性定理要求满足的条件一致，也就是说，若 θ^* 是动态系统的稳定状态，则必须满足条件 $F'(\theta^*)<0$。如果用相位图表示复制动态动力学方程，那就是与 x 轴相交且交点处的切线斜率为负的点为无线传感器网络信任博弈的演化稳定策略。

定理 4-1　若 $G_T+\alpha T-G_D-C>0$，$G_D+C+L-G_T-\alpha T>0$ 且 $2G_T+2\alpha T-2G_D-2C-L>0$，那么 $\theta_1^*=0$ 和 $\theta_2^*=1$ 均是无线传感器网络信任博弈的演化稳定策略且 $p(\theta_1^*=0)<p(\theta_2^*=1)$，其中 $p(\theta_1^*=0)$ 和 $p(\theta_2^*=1)$ 分别表示传感器节点选择动作 Distrust 和 Trust 的概率。

证明　对式(4-4)两边求导得

$$F'(\theta) = -3L\theta^2 + (2G_D + 2C + 4L - 2G_T - 2\alpha T)\theta + G_T + \alpha T - G_D - C - L \tag{4-8}$$

分别令 θ 为 0 和 1 得

$$F'(0) = G_T + \alpha T - G_D - C - L < 0 \tag{4-9}$$

$$F'(1) = G_D + C - G_T - \alpha T < 0 \tag{4-10}$$

由 $2G_T+2\alpha T-2G_D-2C-L>0$ 可得 $G_T+\alpha T-G_D-C>G_D+C+L-G_T-\alpha T$。
所以，可得

$$\begin{aligned} 0 &< (G_D + C + L - G_T - \alpha T)/L \\ &= (G_D + C + L - G_T - \alpha T)/(G_T + \alpha T - G_D - C + G_D + C + L - G_T - \alpha T) \\ &< \frac{G_D + C + L - G_T - \alpha T}{2(G_D + C + L - G_T - \alpha T)} = \frac{1}{2} \end{aligned} \tag{4-11}$$

由式(4-9)至式(4-11)可得复制动态动力学方程式(4-4)的相位图如图 4-1 所示。

由图 4-1 可知，$\theta_1^*=0$ 和 $\theta_2^*=1$ 处切线斜率小于 0，因此 $\theta_1^*=0$ 和 $\theta_2^*=1$ 均是无线传感器网络信任博弈的演化稳定策略。又由式(4-11)可知，参与交互的传感器节点选择动作 Distrust 的概率小于选择动作 Trust 的概率，即 $p(\theta_1^*=0)<p(\theta_2^*=1)$。证毕。

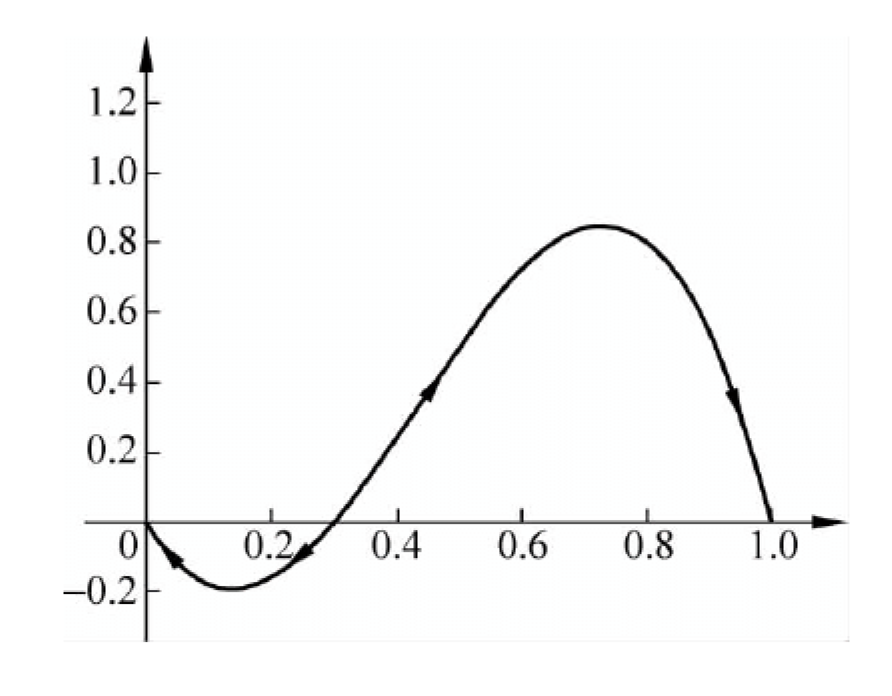


图 4-1　复制动态动力学方程相位图(一)

定理 4-1 代表的含义如下：当参与交互的第一个传感器节点选择动作 Trust 时，由于

$$G_T + G_C + \alpha T - 2C - (G_D + G_C - C) = G_T + \alpha T - G_D - C > 0 \tag{4-12}$$

即参与交互的第二个传感器节点选择动作 Trust 的收益大于选择动作 Distrust 的收益；当参与交互的第一个传感器节点选择动作 Distrust 时，由于

$$G_T + \alpha T - C - L - G_D < 0 \tag{4-13}$$

即参与交互的第二个传感器节点选择动作 Distrust 的收益大于选择动作 Trust 的收益。$\theta_1^* = 0$ 和 $\theta_2^* = 1$ 均是无线传感器网络信任博弈的演化稳定策略，意味着动作 Trust 和 Distrust 都有可能被参与交互的传感器节点选择。

定理 4-2 若 $G_T + \alpha T - G_D - C > 0$，$G_D + C + L - G_T - \alpha T > 0$ 且 $2G_T + 2\alpha T - 2G_D - 2C - L < 0$，则 $\theta_1^* = 0$ 和 $\theta_2^* = 1$ 均是无线传感器网络信任博弈的演化稳定策略且 $p(\theta_1^* = 0) > p(\theta_2^* = 1)$。

证明 与定理 4-1 的证明过程类似，可得

$$F'(0) = G_T + \alpha T - G_D - C - L < 0 \tag{4-14}$$

$$F'(1) = G_D + C - G_T - \alpha T < 0 \tag{4-15}$$

$$\frac{1}{2} < (G_D + C + L - G_T - \alpha T)/L < 1 \tag{4-16}$$

由式(4-14)至式(4-16)可得传感器节点信任演化的复制动态动力学方程式(4-4)的相位图如图 4-2 所示。

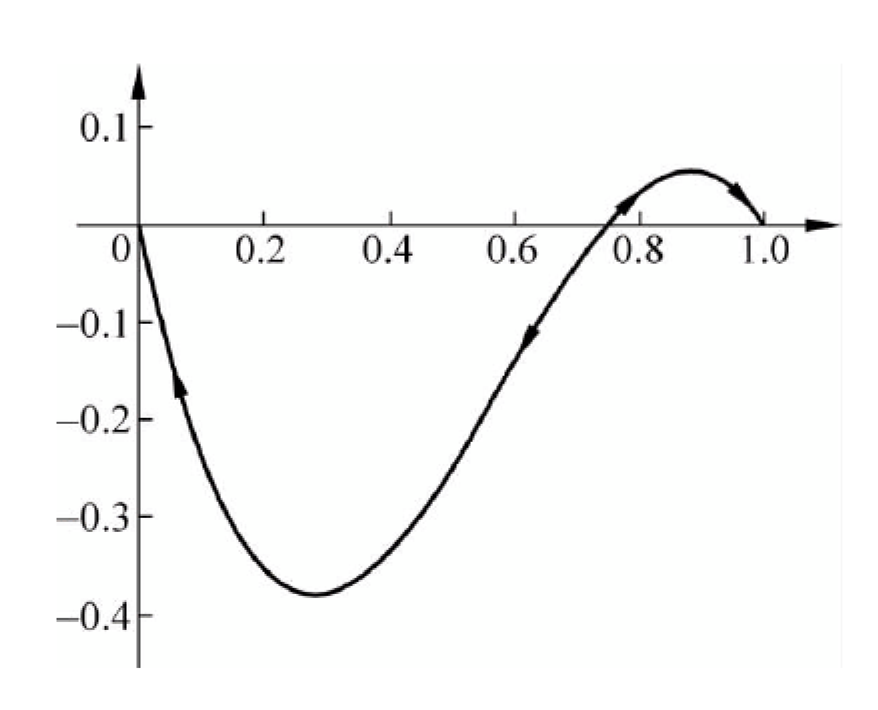


图 4-2 复制动态动力学方程相位图(二)

由图 4-2 可知，$\theta_1^* = 0$ 和 $\theta_2^* = 1$ 处切线斜率小于 0，因此 $\theta_1^* = 0$ 和 $\theta_2^* = 1$ 均是无线传感器网络信任博弈的演化稳定策略。又由式(4-16)可知，参与交互的传感器节点选择动作 Distrust 的概率大于选择动作 Trust 的概率，即 $p(\theta_1^* = 0) > p(\theta_2^* = 1)$。证毕。

定理 4-3 若 $G_T + \alpha T - G_D - C < 0$，则 $\theta_1^* = 0$ 是无线传感器网络信任博弈的演化稳定策略。

证明 易得

$$F'(0) = G_T + \alpha T - G_D - C - L < 0 \tag{4-17}$$

$$F'(1) = G_D + C - G_T - \alpha T > 0 \tag{4-18}$$

由式(4-17)、式(4-18)可得传感器节点信任演化的复制动态动力学方程式(4-4)的相位图如图 4-3 所示。

由图 4-3 可知，只有 $\theta_1^* = 0$ 处切线斜率小于 0，因此只有 $\theta_1^* = 0$ 是无线传感器网络信任

博弈的演化稳定策略。证毕。

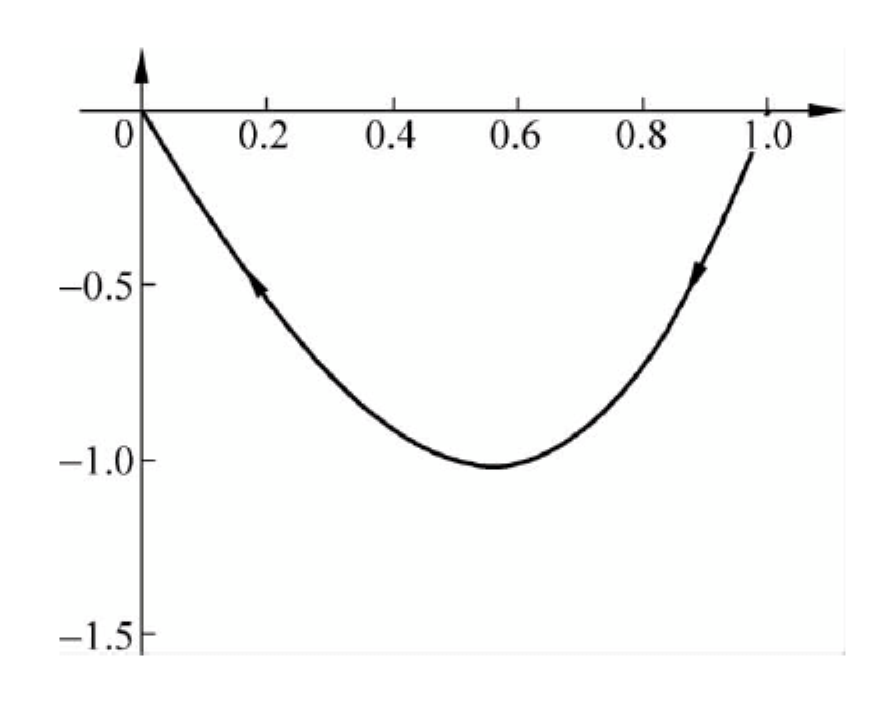


图 4-3 复制动态动力学方程相位图(三)

定理 4-3 表明,无论参与交互的第一个传感器节点选择动作是 Distrust 还是 Trust,参与交互的第二个传感器节点选择动作 Trust 的收益总是小于选择动作 Distrust 的收益。最终选择动作 Trust 的参与交互的传感器节点数比例会稳定在 $\theta_1^*=0$ 处,即都选择动作 Distrust,这将导致整个无线传感器网络中的传感器节点都处在互相不合作的状态。

定理 4-4 若 $G_T+\alpha T-G_D-C-L>0$,则 $\theta_2^*=1$ 是无线传感器网络信任博弈的演化稳定策略。

证明 易得

$$F'(0)=G_T+\alpha T-G_D-C-L>0 \tag{4-19}$$

$$F'(1)=G_D+C-G_T-\alpha T<G_D+C+L-G_T-\alpha T<0 \tag{4-20}$$

由式(4-19)、式(4-20)可得传感器节点信任演化的复制动态动力学方程式(4-4)的相位图如图 4-4 所示。

由图 4-4 可知,只有 $\theta_2^*=1$ 处切线斜率小于 0,因此只有 $\theta_2^*=1$ 是无线传感器网络信任博弈的演化稳定策略。证毕。

定理 4-4 表明,无论参与交互的第一个传感器节点选择动作是 Distrust 还是 Trust,参与交互的第二个传感器节点选择动作 Trust 的收益总是大于选择动作 Distrust 的收益。最终选择动作 Trust 的参与交互的传感器节点数比例会稳定在 $\theta_2^*=1$ 处,即都选择动作 Trust。实际上,当定理 4-4 的条件满足时,动作 Trust 已成为无线传感器网络信任博弈的严格占优策略。

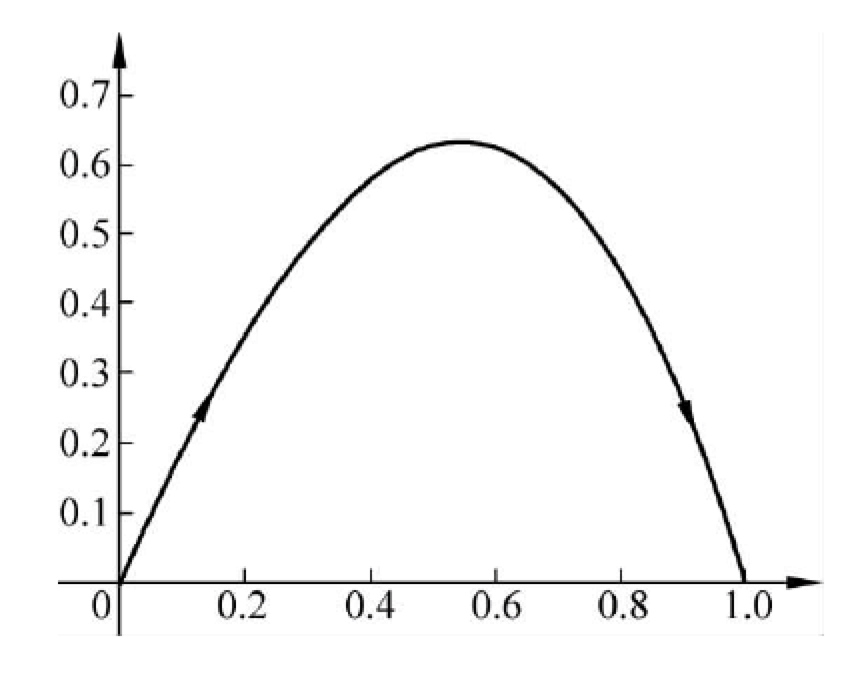


图 4-4 复制动态动力学方程相位图(四)

由定理 4-1 至定理 4-4 可知，要促使参与交互的传感器节点选择动作 Trust，从而保证无线传感器网络的安全性和稳定性，设计的信任管理机制应满足定理 4-1 或定理 4-4 的条件。引入的 αT 使传感器节点在信任博弈的过程中增加了激励机制，当定理 4-1 条件满足并逐步增大 αT 时，$\theta_3^*=(G_D+C+L-G_T-\alpha T)/L\rightarrow 0$，这意味着随着无线传感器网络信任博弈的进行，选择动作 Distrust 的参与交互的传感器节点数比例将逐渐降低，最后达到一个稳定的低比例水平。当 αT 增大到一定程度使得定理 4-4 的条件满足时，整个无线传感器网络将处于理想的稳定状态，此时无论参与交互的传感器节点开始选择何种策略，最终都将选择动作 Trust 作为稳定状态。对于定理 4-2 和定理 4-3 满足的条件是设计无线传感器网络信任管理机制时必须要避免的情况，因为它们意味着参与交互的传感器节点选择动作 Distrust 的概率大于选择动作 Trust 的概率或将选择动作 Distrust 作为最终的稳定状态，这会导致无线传感器网络处于不稳定状态。

4.4 实验

实验环境使用 MATLAB R2010a，通过设置 G_T、G_D、C、L、αT 不同的取值，来验证无线传感器网络信任博弈中的演化稳定策略和激励机制所起的作用。实验分成两组：第一组使得设置的博弈参数取值分别满足定理 4-1 至定理 4-4 的条件，再观察无线传感器网络传感器节点信任演化曲线的变化情况；第二组通过改变 αT 值，再观察激励机制在无线传感器网络传感器节点信任演化过程中起到的作用。

4.4.1 演化稳定策略定理的数值验证

为满足定理 4-1 至定理 4-4 的条件，分别设定：①$G_T=11$，$G_D=3$，$C=8$，$\alpha T=3$，$L=5$；②$G_T=11$，$G_D=5$，$C=8$，$\alpha T=3$，$L=5$；③$G_T=9$，$G_D=5$，$C=8$，$\alpha T=3$，$L=5$；④$G_T=13$，$G_D=3$，$C=8$，$\alpha T=3$，$L=4$。图 4-5 至图 4-8 分别给出了 4 种情况下无线传感器网络传感器节点信任演化的变化曲线。

图 4-5 中博弈参数的取值满足定理 4-1 的条件。从图中可以看出，当传感器节点信任演化的复制动态动力学方程式(4-4)的初始值为 0.401，即 40.1%的无线传感器网络传感器节点选择动作 Trust 时，参与交互的传感器节点通过试错和模仿，不断调整自己的策略，约经过 38 次博弈，最终选择动作 Trust 的参与交互的传感器节点数比例稳定在 $\theta_2^*=1$ 处。这意味着初始无线传感器网络传感器节点选择动作 Trust 的比例数只要大于 40.1%，则参与交互的传感器节点最终都会选择动作 Trust。当传感器节点信任演化的复制动态动力学方程式(4-4)的初始值为 0.399，即 39.9%的无线传感器网络传感器节点选择动作 Trust 时，约经过 44 次博弈，最终选择动作 Trust 的参与交互的传感器节点数比例稳定在 $\theta_1^*=0$。这意味着初始无线传感器网络传感器节点选择动作 Trust 的比例数只有小于 39.9%，参与交互的传感器节点最终才会选择动作 Distrust。这些实验结果反映出 $\theta_1^*=0$ 和 $\theta_2^*=1$ 均是无线传感器网络信任博弈的演化稳定策略且 $p(\theta_1^*=0)<p(\theta_2^*=1)$。

图 4-6 中博弈参数的取值满足定理 4-2 的条件。从图中可以看出，当传感器节点信任演化的复制动态动力学方程式(4-4)的初始值为 0.801，即 80.1%的无线传感器网络传感器节点选择动作 Trust 时，约经过 30 次博弈，最终选择动作 Trust 的参与交互的传感器节点

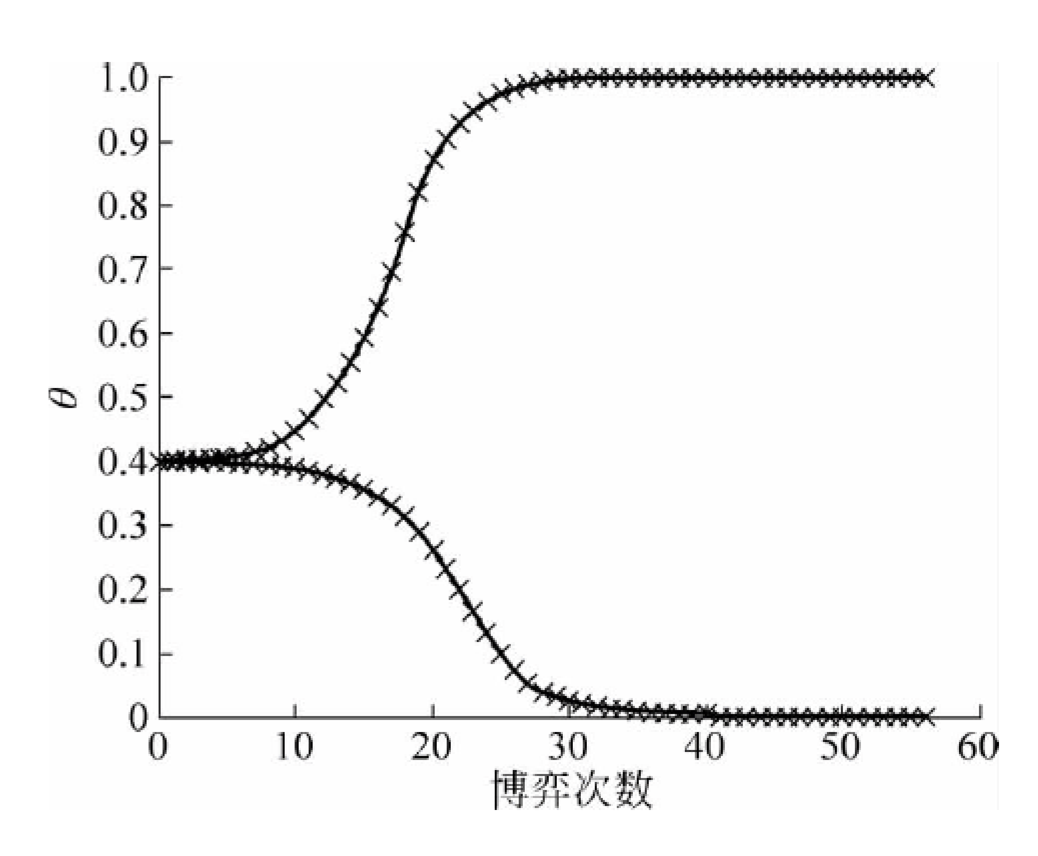


图 4-5　传感器节点信任演化曲线(一)

数比例稳定在 $\theta_2^*=1$ 处。当传感器节点信任演化的复制动态动力学方程式(4-4)的初始值为 0.799，即 79.9%的参与交互的传感器节点选择动作 Trust 时，约经过 56 次博弈，最终选择动作 Trust 的参与交互的传感器节点数比例稳定在 $\theta_1^*=0$。这些实验结果反映出 $\theta_1^*=0$ 和 $\theta_2^*=1$ 均是无线传感器网络信任博弈的演化稳定策略且 $p(\theta_1^*=0)>p(\theta_2^*=1)$。

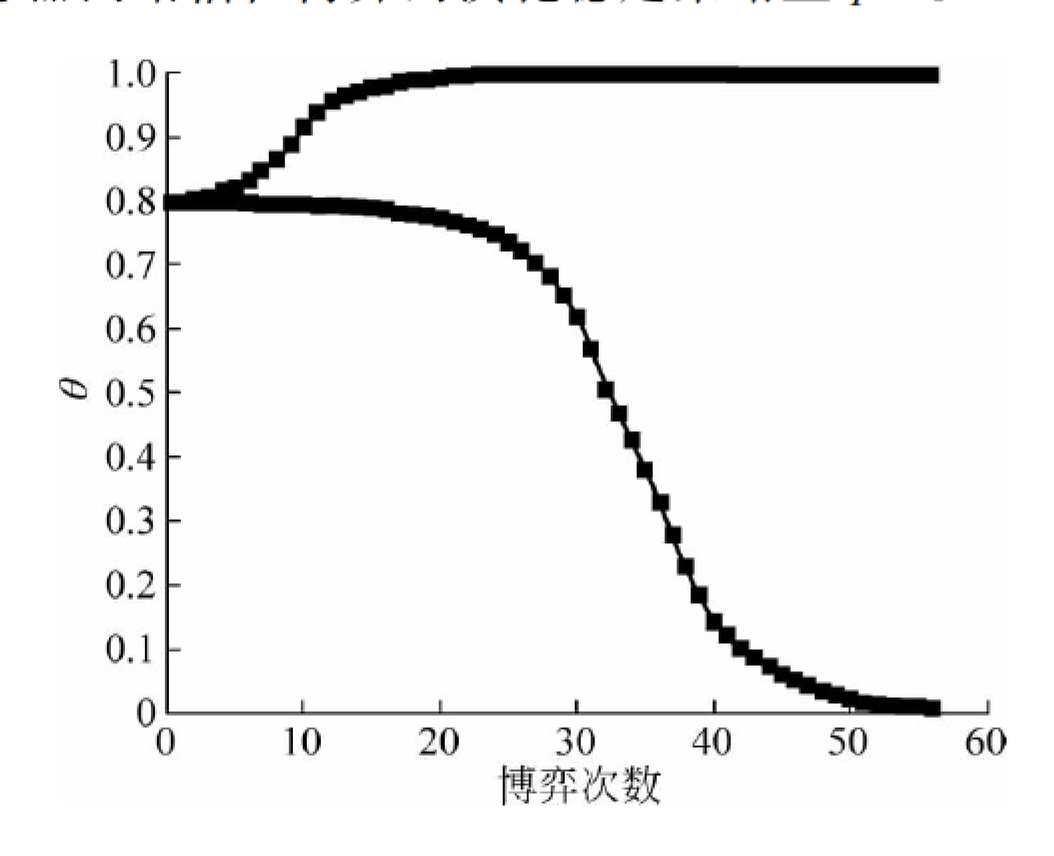


图 4-6　传感器节点信任演化曲线(二)

图 4-7 中博弈参数的取值满足定理 4-3 的条件。从图中可看出，即使将传感器节点信任演化的复制动态动力学方程式(4-4)的初始值设定为 0.999，即 99.9%的无线传感器网络传感器节点初始选择动作 Trust，但约经过 58 次博弈，最终选择动作 Trust 的参与交互的传感器节点数比例稳定在 $\theta_1^*=0$。实验结果反映出 $\theta_1^*=0$ 是无线传感器网络信任博弈的演化稳定策略。

图 4-8 中博弈参数的取值满足定理 4-4 的条件。从图中可看出，只要将传感器节点信任演化的复制动态动力学方程式(4-4)的初始值设定为 0.001，即 0.1%的无线传感器网络传感器节点初始选择动作 Trust，约经过 53 次博弈，最终选择动作 Trust 的参与交互的传感器节点数比例就能稳定在 $\theta_2^*=1$。实验结果反映出 $\theta_2^*=1$ 是无线传感器网络信任博弈的演化稳定策略。

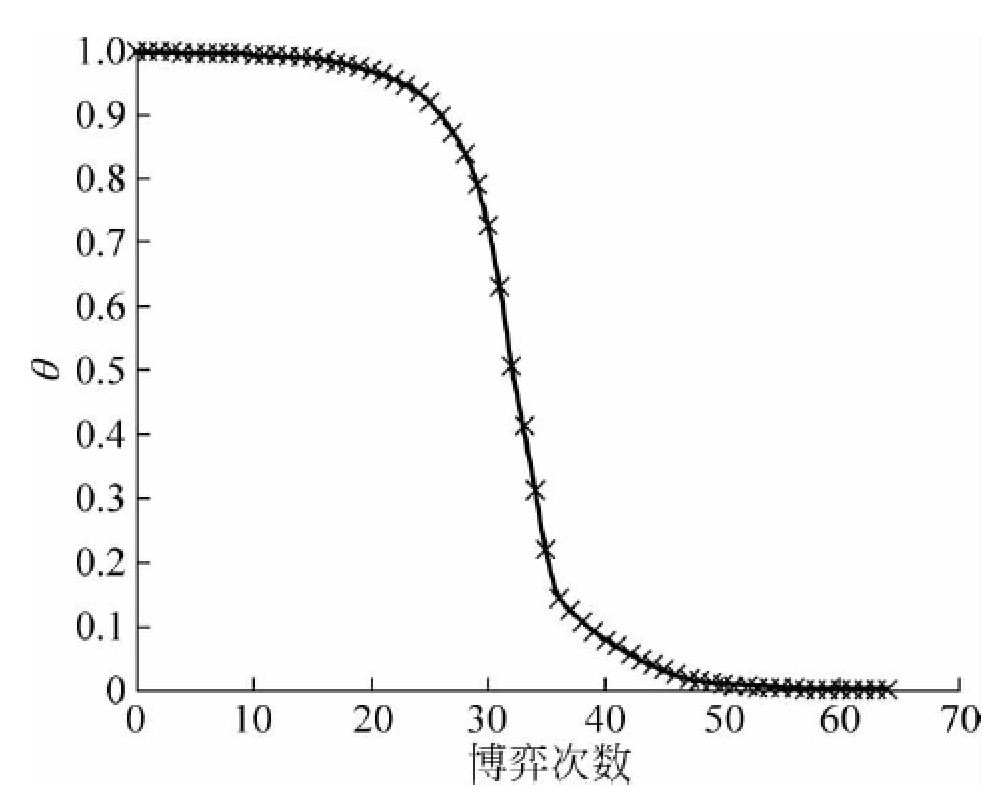


图 4-7　传感器节点信任演化曲线(三)

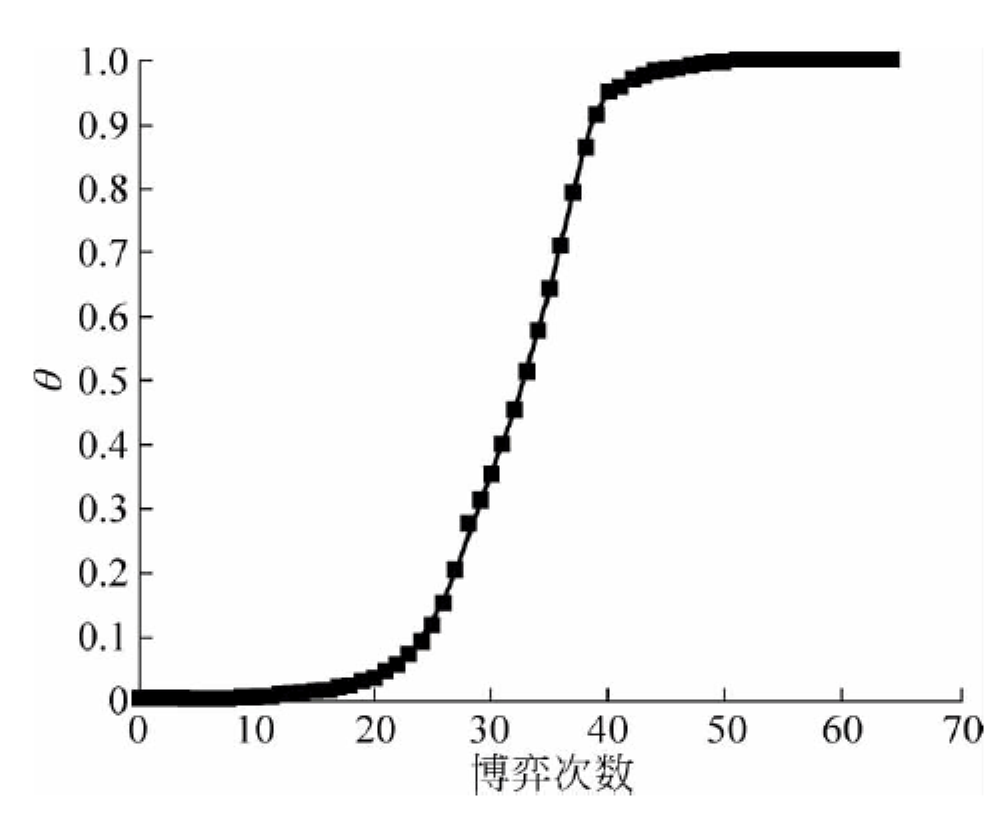


图 4-8　传感器节点信任演化曲线(四)

4.4.2　激励机制的效果

为了说明激励机制在促进传感器节点选择动作 Trust 上所起的作用,分别设定:①$G_T=11$, $G_D=3$, $C=8$, $\alpha T=3$, $L=5$;②$G_T=11$, $G_D=3$, $C=8$, $\alpha T=3.5$, $L=5$。图 4-9 给出了无线传感器网络传感器节点不同初始选择动作 Trust 比例数下的信任演化曲线。图 4-10 给出了无线传感器网络传感器节点相同初始选择动作 Trust 比例数下信任向 $\theta_2^*=1$ 演化的变化曲线。

从图 4-9 可以看出,当 $\alpha T=3$ 时,无线传感器网络传感器节点信任演化的临界值为 0.401,而当 $\alpha T=3.5$ 时,临界值变为 0.301。这意味着当 αT 值从 3 增加到 3.5 后,即使参与交互的传感器节点初始选择动作 Trust 比例数由 40.1%降为 30.1%,但随着博弈的进行,最终 $\theta_2^*=1$ 仍将作为无线传感器网络信任博弈的演化稳定策略。

从图 4-10 中可以看出,在传感器节点信任演化的复制动态动力学方程式(4-4)设置相同初始值 0.401 情况下,当 $\alpha T=3$ 时,要经过约 38 次博弈,而当 $\alpha T=3.5$ 时,只需经过约 20 次博弈即能达到系统稳定点 $\theta_2^*=1$。很显然,$\alpha T=3.5$ 对应的传感器节点信任演化曲线收敛到系统稳定状态的速度明显快于 $\alpha T=3$ 对应的曲线。

图 4-9 和图 4-10 的实验结果反映出激励机制在促进传感器节点选择动作 Trust 上明显起作用,也就是说,通过采用与信任度绑定的激励机制,奖励信任合作,将有利于无线传感

器网络向选择动作 Trust(即信任策略)的稳定状态演化,从而保障无线传感器网络的稳定和安全。

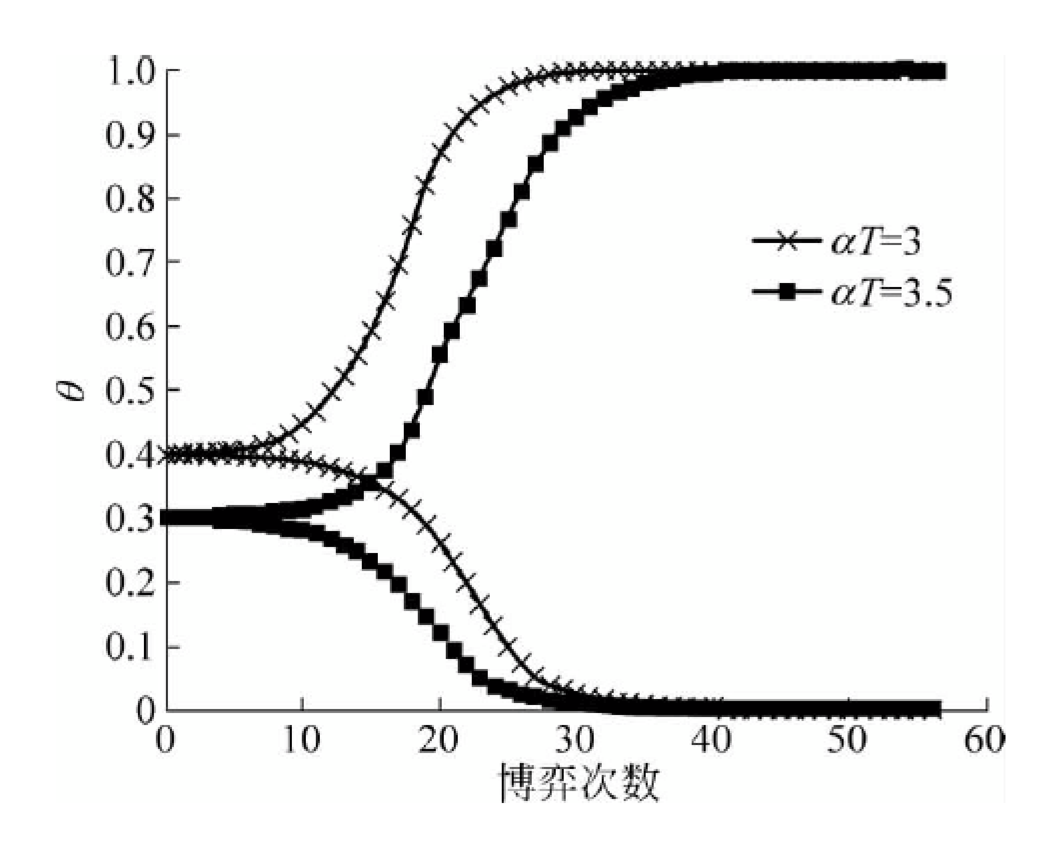


图 4-9 激励机制下的传感器节点信任演化曲线(一)

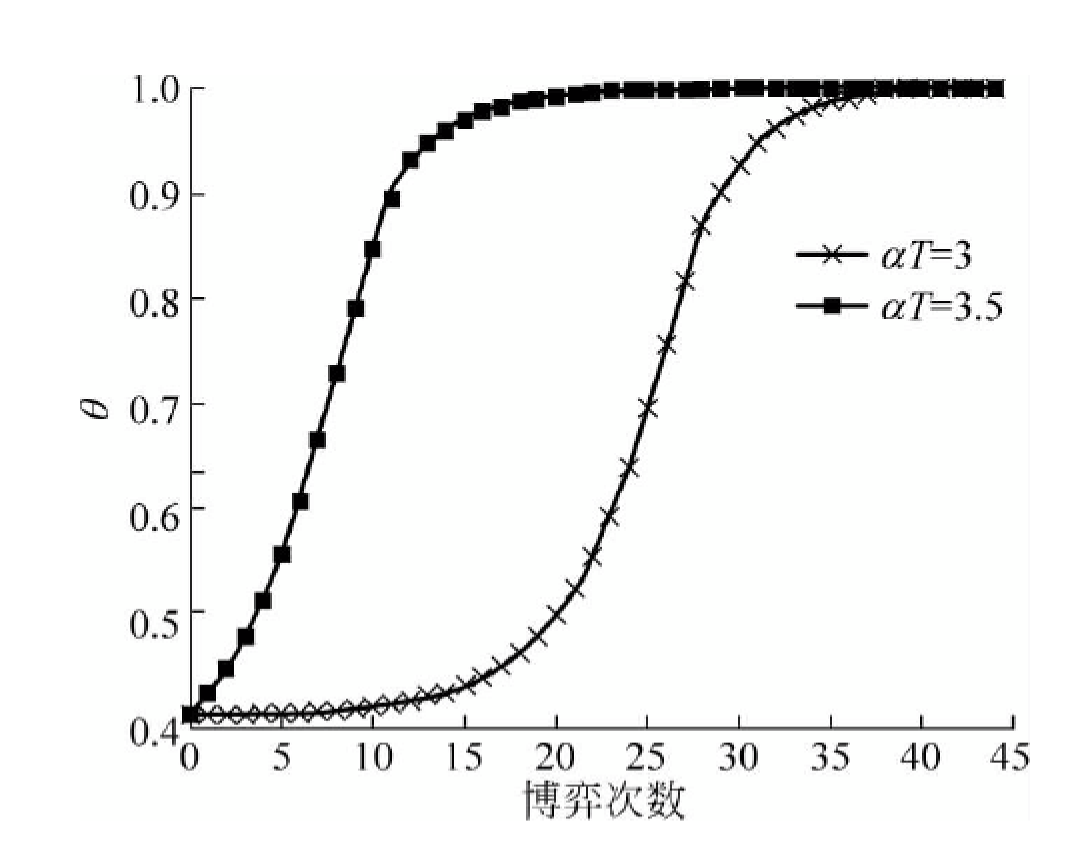


图 4-10 激励机制下的传感器节点信任演化曲线(二)

4.5 小结

基于信任的安全机制是无线传感器网络中重要的安全技术之一。这种无线传感器网络传感器节点间的信任关系能帮助各传感器节点建立信心,降低合作风险,从而保证整个无线传感器网络的安全和稳定。本章利用演化博弈针对参与交互的传感器节点的信任决策过程所建立的模型反映了参与交互的传感器节点在选择不同策略时的收益得失,能体现无线传感器网络传感器节点行为的有限理性、博弈的非零和性和重复性,以及参与交互的传感器节点选择动作 Trust 的模仿性。与信任度绑定的激励机制有效地降低了无线传感器网络传感器节点初始选择动作 Trust 比例数的要求,提高了系统向信任稳定状态演化的收敛速度。给出的无线传感器网络信任博弈的复制动态动力学方程为寻找种群的演化稳定策略奠定了基础,得到了在不同参数条件下,参与交互的传感器节点如何通过不断地模仿与试错来动态调整自己的信任决策,最终达到系统的稳定,从而为无线传感器网络信任机制的设计提供了理论基础。

第 5 章

基于微分博弈的无线传感器网络恶意程序传播机制研究

本章扩展经典流行病理论使之适合无线传感器网络恶意程序传播现状，并引入不同的参数来揭示无线传感器网络恶意程序传播过程。然后将恶意程序在无线传感器网络传播时“无线传感器网络系统”和“恶意程序”之间的决策问题看作优化控制问题，建立相应的微分博弈模型，在“恶意程序”动态改变其策略的前提下，得到“无线传感器网络系统”的最优控制策略，为无线传感器网络恶意程序传播的防御机制的设计奠定理论基础。

5.1 引言

近期研究[171, 172]表明，恶意程序(Malware)在由大量传感器节点组成的无线传感器网络中容易传播流行。例如，Yang 等人[172]通过在 Mica2 节点上的实验，说明了在无线传感器网络中传播传感器节点蠕虫(Sensor Worm)的方便性。实际上，这种恶意程序的易传播性是与无线传感器网络的特点密切相关的。一方面，由于传感器节点资源的限制使得它们没有足够的能力保护自身的系统安全，与传统的计算机系统相比，它们更容易因恶意程序的攻击而被捕获。在同一个无线传感器网络中的传感器节点一般在其硬件和软件方面都具有同构性，一旦有一种恶意程序攻陷一个传感器节点，这种同构性将使得所有的传感器节点易于被同一种恶意程序攻陷。另一方面，当前一些不需要物理连接只需要通过空气传播的重编程协议(Reprogramming Protocol)已经被提出并用于传感器节点的重编程(Reprogramming)和重配置(Reconfigure)，典型的主要有 Trickle、Firecracker、Deluge 和 MNP[173]。重编程协议能为传感器节点提供系统软件分发和更新功能，尤其对那些部署后不能再进行人工操作的节点而言，这种重编程技术对它们来说是一项必不可少的基础服务。然而，这种技术同时也给恶意程序在整个无线传感器网络中的传播提供了一条途径。

恶意程序一旦利用传感器节点的系统软件漏洞在无线传感器网络中广泛传播后，它们就能窃听传感器节点感知的数据，甚至可以采用耗尽传感器节点能量的方法使传感器节点完全处于瘫痪状态，从而严重影响整个无线传感器网络数据的机密性和整个网络工作的稳定性。因此，本章主要研究以下两个关键问题：

（1）无线传感器网络中的恶意程序是如何传播的？

（2）如何控制无线传感器网络中的恶意程序传播？什么样的策略是控制恶意程序传播

的最优控制策略?

为了得到控制无线传感器网络中恶意程序传播的最优策略,首先应考虑建立无线传感器网络恶意程序的传播模型。因为流行病理论(Epidemic Theory)已成功应用于传统计算机网络中的蠕虫传播[174]、错误传递(Fault Propagation)[175]和信息扩散中的流行病算法(Epidemic Algorithm)[176,177]等领域的建模,所以本章考虑利用流行病理论解决无线传感器网络中恶意程序传播的建模问题。实际上,流行病理论早期主要用于研究人类社会中的传染病传播问题,利用数学公式明确地表达病因、宿主和环境之间的传染病流行规律,从理论上探讨如何控制传染病的传播问题[178]。在流行病理论中,根据各种传染病的不同情况,每个个体都对应到某一状态。这些状态主要有易感(Susceptible, S)、被感染(Infected, I)、潜伏(Exposed, E)、康复(Recovered, R)等,其中状态 S 表示个体处于当前未被病毒感染但属于易感的状态;状态 I 表示个体当前已被病毒感染并能将病毒传染给别的个体的状态;状态 R 表示被病毒感染的个体经过治疗或易感个体经过预防后对同一种病毒已具有免疫力的状态;状态 E 表示个体当前已被病毒感染但无病毒传播能力的状态。相应地,根据不同传染病的特性,研究者组合上述的不同状态,得到了不同的流行病模型 SI[179,180]、SIR[181,182]、SEIR[183,184]等。由于网络环境下的信息传播与传染病传播有类似的特性,所以可以考虑将原先适用于流行病传播领域的模型借鉴到不同的网络环境中。

在恶意程序传播模型建立的同时还需考虑无线传感器网络攻防双方的决策问题。对无线传感器网络这个系统而言,安装安全补丁是一种通用地保障整个网络系统安全运行的方法,这种方法一方面可通过修复系统漏洞阻止容易感染恶意程序的传感器节点被感染,另一方面也能治愈被恶意程序感染的传感器节点并使它对同一种恶意程序具有免疫力。无线传感器网络系统也可以采取让已感染恶意程序的传感器节点强制进入休眠状态,从而避免恶意程序的传播。然而,这些方法都不可避免地会对无线传感器网络的正常运行带来干扰。例如,安装安全补丁的过程将消耗无线传感器网络本来就低的带宽从而影响正常数据的传输;过多的传感器节点强制进入休眠状态将影响整个网络的数据通信。因此,无线传感器网络系统面临的一个挑战就是如何在保证正常通信和最小化恶意程序影响之间选择最优的控制策略。对恶意程序而言,它可以传染更多的易感传感器节点从而可以窃听更多由传感器节点感知的数据信息,也可以直接使被感染的传感器节点失去正常功能从而给整个无线传感器网络造成瞬时的高损失。但是,传染太多的易感传感器节点,容易被无线传感器网络系统发现导致其采取预防措施,恶意程序也不能从已失去功能的传感器节点上获取需要的信息或失去利用被感染的传感器节点传播恶意程序的机会。因此,恶意程序也面临在何时传播自己和是否使被感染传感器节点失去功能方面的最优控制策略选择问题。

博弈论是解决优化策略问题的有效方法之一,根据不同博弈环境常需要考虑不同的博弈类型。其中微分博弈通常关注连续时间的决策演化过程,并采用微分等式描述系统的连续动态变化,这些特点与恶意程序传播过程中表现出的特性完全一致,因此这种动态博弈类型非常适合用于说明恶意程序传播过程中的收益得失。

本章首先扩展经典流行病理论使之适合无线传感器网络恶意程序传播现状,并引入不同的参数来揭示无线传感器网络恶意程序传播过程。其中,对经典流行病理论的扩展主要根据无线传感器网络的特性展开,这些特性包括:传感器节点为节省能量消耗而需要周期性地进入休眠状态,那些被恶意程序感染的传感器节点在进入休眠状态后不能将恶意程序

传播到其他传感器节点；而处于任何状态的传感器节点都会在它们的能量消耗殆尽时失去所有功能。然后本章将恶意程序在无线传感器传播时“无线传感器网络系统”和“恶意程序”之间的决策问题看作优化控制问题，并利用微分博弈为“无线传感器网络系统”得到最优控制策略，这种策略将在考虑“恶意程序”最大化破坏无线传感器网络的前提下，最小化“无线传感器网络系统”和“恶意程序”产生的成本。

在扩展作者前期工作[185]的基础上，本章的工作主要包括以下内容：

(1) 通过扩展经典流行病模型得到一种新的流行病传播模型，这种模型能准确地反映出传感器节点因节省能量消耗而周期性地进入休眠状态以及传感器节点在能量耗尽时将失去功能等特性。

(2) 建立一个“无线传感器网络系统”和“恶意程序”之间的零和“恶意程序防御微分博弈”模型，该模型能体现“无线传感器网络系统”和“恶意程序”双方在恶意程序传播时的交互情况，能反映它们之间的收益得失。

(3) 在考虑“恶意程序”最大化破坏无线传感器网络的前提下，为“无线传感器网络系统”得到了最优的控制策略，这些策略能明显地抑制恶意程序在无线传感器网络中的传播，同时因计算方便而易于在资源有限的传感器节点上实现。

本章其余章节安排如下：5.2 节介绍相关工作；5.3 节讨论要研究的无线传感器网络环境并通过扩展经典流行病模型得到一种适合描述无线传感器网络中恶意程序传播的流行病模型；5.4 节给出恶意程序防御微分博弈模型的定义，并描述如何得到无线传感器网络系统的最优控制策略；5.5 节通过实验揭示无线传感器网络恶意程序流行病模型中各状态的动态演化过程，并说明“无线传感器网络系统”和“恶意程序”采取的控制策略对无线传感器网络恶意程序传播的影响，此外，还验证了无线传感器网络系统最优控制策略的有效性；5.6 节给出本章小结。

本章涉及的符号含义如下：

N 表示整个无线传感器网络的传感器节点数。

σ 表示传感器节点在撒播区域上的分布密度。

r 表示传感器节点信号的最大发射距离。

S 表示传感器节点存在被恶意程序发现的硬件或系统漏洞，处于易被恶意程序感染的状态。

$\widetilde{S}$ 表示传感器节点存在被恶意程序发现的硬件或系统漏洞但正在休眠的状态。

I 表示传感器节点处于已被恶意程序感染的状态。

$\widetilde{I}$ 表示传感器节点处于已被恶意程序感染且正在休眠的状态。

R 表示传感器节点处于对恶意程序有免疫力的状态。

$\widetilde{R}$ 表示传感器节点处于对恶意程序有免疫力且正在休眠的状态。

D 表示传感器节点处于完全失去所有功能的状态。

I_0 表示初始已感染恶意程序节点数。

$S(t)$ 表示传感器节点在时刻 t 处于状态 S 的数量。

$\widetilde{S}(t)$ 表示传感器节点在时刻 t 处于状态 $\widetilde{S}$ 的数量。

$I(t)$ 表示传感器节点在时刻 t 处于状态 I 的数量。

$\widetilde{I}(t)$表示传感器节点在时刻 t 处于状态$\widetilde{I}$ 的数量。

$R(t)$表示传感器节点在时刻 t 处于状态 R 的数量。

$\widetilde{R}(t)$表示传感器节点在时刻 t 处于状态$\widetilde{R}$ 的数量。

$D(t)$表示传感器节点在时刻 t 处于状态 D 的数量。

γ_{SI} 表示传感器节点从状态 S 转换到 I 的概率。

$\gamma_{S\widetilde{S}}$ 表示传感器节点从状态 S 转换到$\widetilde{S}$ 的概率。

γ_{SR} 表示传感器节点从状态 S 转换到 R 的概率。

γ_{SD} 表示传感器节点从状态 S 转换到 D 的概率。

$\gamma_{I\widetilde{I}}$ 表示传感器节点从状态 I 转换到$\widetilde{I}$ 的概率。

γ_{IR} 表示传感器节点从状态 I 转换到 R 的概率。

γ_{ID} 表示传感器节点从状态 I 转换到 D 的概率。

$\gamma_{R\widetilde{R}}$ 表示传感器节点从状态 R 转换到$\widetilde{R}$ 的概率。

γ_{RD} 表示传感器节点从状态 R 转换到 D 的概率。

$\gamma_{\widetilde{S}S}$表示传感器节点从状态$\widetilde{S}$ 转换到 S 的概率。

$\gamma_{\widetilde{I}I}$表示传感器节点从状态$\widetilde{I}$ 转换到 I 的概率。

$\gamma_{\widetilde{R}R}$表示传感器节点从状态$\widetilde{R}$ 转换到 R 的概率。

$\gamma_{SR}^{S}(t)$表示无线传感器网络系统在时刻 t 将状态 S 的传感器节点转换为状态 R 的主观努力程度。

$\gamma_{IR}^{S}(t)$表示无线传感器网络系统在时刻 t 将状态 I 的传感器节点转换为状态 R 的主观努力程度。

$\gamma_{I\widetilde{I}}^{S}(t)$表示无线传感器网络系统在时刻 t 将状态 I 的传感器节点转换为状态$\widetilde{I}$ 的主观努力程度。

$\gamma_{SI}^{M}(t)$表示恶意程序在时刻 t 将状态 S 的传感器节点转换为状态 I 的主观努力程度。

$\gamma_{ID}^{M}(t)$表示恶意程序在时刻 t 将状态 I 的传感器节点转换为状态 D 的主观努力程度。

$\mathbb{G}$ 表示本章定义的无线传感器网络恶意程序防御微分博弈。

$\mathcal{N}$表示博弈的参与者集合。

$\mathcal{U}$表示参与者无线传感器网络系统可用的控制策略集合。

$\mathcal{V}$表示参与者恶意程序可用的控制策略集合。

$\mathcal{C}$表示两个参与者控制策略集合的笛卡儿积。

$f(t,\boldsymbol{x}(t),\mu(t),\nu(t))$表示状态函数。

$\boldsymbol{x}(t)$表示七维的状态向量。

$\mathcal{F}$表示状态函数集合。

$J(\mu(t),\nu(t))$表示反映无线传感器网络 QoS 的成本函数。

$g(t,\boldsymbol{x}(t),\mu(t),\nu(t))$表示在时刻 t 的瞬时成本。

$q(\boldsymbol{x}(T))$表示逗留期 T 结束时的终期成本。

c_I 表示被感染传感器节点产生的瞬时成本对应的系数。

c_D 表示被感染传感器节点因被恶意程序故意杀死产生的瞬时成本对应的系数。

c_D^T 表示终期成本对应的系数。

c_{IR} 表示将被感染传感器节点修复后产生的收益对应的系数。

c_R 表示安装安全补丁过程中产生的瞬时成本对应的系数。

$c_{I\tilde{I}}$ 表示将被感染传感器节点转换为休眠状态后产生的收益对应的系数。

$c_{\tilde{S}}$ 表示将易感传感器节点转换为休眠状态后产生的瞬时成本对应的系数。

$c_{\tilde{R}}$ 表示将康复传感器节点转换为休眠状态后产生的瞬时成本对应的系数。

$c_{\tilde{I}}$ 表示隔离被感染传感器节点产生的收益对应的系数。

$\mu^*(t)$表示无线传感器网络系统的最优控制策略。

$\nu^*(t)$表示恶意程序的最优控制策略。

V 表示无线传感器网络恶意程序防御微分博弈达到鞍点时的价值函数值。

$\boldsymbol{x}^*(t)$表示无线传感器网络恶意程序防御微分博弈达到鞍点时最优的状态轨迹(State Trajectory)。

$S^*(t)$表示无线传感器网络恶意程序防御微分博弈达到鞍点时易感传感器节点的数量。

$I^*(t)$表示无线传感器网络恶意程序防御微分博弈达到鞍点时被感染传感器节点的数量。

$\gamma_{IR}^{S*}(t)$表示无线传感器网络恶意程序防御微分博弈达到鞍点时无线传感器网络系统将传感器节点从状态 I 转换为 R 的最优主观努力程度。

$\gamma_{I\tilde{I}}^{S*}(t)$表示无线传感器网络恶意程序防御微分博弈达到鞍点时无线传感器网络系统将传感器节点从状态 I 转换为$\tilde{I}$ 的最优主观努力程度。

$\gamma_{SI}^{M*}(t)$表示无线传感器网络恶意程序防御微分博弈达到鞍点时恶意程序将传感器节点从状态 S 转换为 I 的最优主观努力程度。

$\gamma_{ID}^{M*}(t)$表示无线传感器网络恶意程序防御微分博弈达到鞍点时恶意程序将传感器节点从状态 I 转换为 D 的最优主观努力程度。

$\boldsymbol{p}(t)$表示协状态函数(Co-state Function)向量。

$H(t,\boldsymbol{p}(t),\boldsymbol{x}(t),\mu(t),\nu(t))$表示哈密尔顿(Hamiltonian)等式。

5.2 相关工作

近年来,伴随着计算机网络的发展,恶意程序已经成为威胁网络安全的主要因素之一。面对恶意代码数量的日益庞大及其威胁的日益严重,由于技术的局限性,仍有大量恶意代码无法被有效监测使恶意代码的防范形势变得日益严峻[186]。要防御恶意程序,首先要获得恶意程序的特征(Signature),从而为获得未知恶意程序特征和已知恶意程序的变种奠定基础。因此,恶意程序检测技术一直是研究者的热门领域。

Egele 等人[187]综述了用于采集潜在恶意程序的动态分析技术及相应的辅助分析程序。Santos 等人[188]提出一种基于操作码序列(Opcode Sequence)的分类器用于检测未知的恶意程序代码。苗甫等人[189]提出一种基于流量统计指纹的恶意代码检测模型,通过提取网络

流量中的包层特征和流层特征，利用两类特征的概率密度函数建立恶意代码流量统计指纹，实现网络中恶意代码流量的检测。孔德光等人[190]从操作码分布序列、调用流图特征、系统调用序列图这3个特征维度归纳和分析恶意代码特征，提出了一种基于多维特征的迷惑恶意代码检测算法。王蕊等人[191]采用可回溯的动态污点分析方法，对恶意代码变种进行细粒度地分析，挖掘其行为特征，从而得到一种抗混淆的恶意代码变种检测方法。张鹏涛等人[192]从指令频率和包含相应指令的文件频率两个方面出发，提出了一种基于惩罚因子的阴性选择算法实现恶意程序的检测。Dube等人[193]结合决策树机器学习算法和静态启发法(Static Heuristics)提出了一种恶意程序标识(Malware Target Recognition)系统。Chen等人[194]为恶意程序的监督学习提出了一种新的分类框架，该框架利用支持向量机和决策树构建学习模型，使用自组织映射实现分类。Perdisci等人[195]提出一种适用于HTTP恶意程序聚类的方法，通过轮流检查网络的命令控制和网络外围通信能自动地得到恶意程序的特征。Chandramohan和Tan[196]关注智能手机的移动恶意程序的检测问题。李鹏和王汝传[197]将自相似特性技术引入到恶意代码的动态分析中，通过跟踪同类型的恶意程序，提取恶意程序的关键特征信息，得到时间调用序列，实现了同种恶意程序的检测。他们[198]还利用基于未知恶意代码样本空间关系特征的自动检测技术，划分恶意代码样本空间关系区域，提取恶意程序特征向量，建立空间关系特征向量索引实现未知恶意代码的检测。王蕊等人[186]结合指令层的污点传播分析与行为层的语义分析，提取出恶意代码的关键行为，利用抗混淆引擎识别语义，得到恶意代码行为特征，从而实现恶意代码行为的检测。其他的检测方法还有基于行为特征的恶意代码检测模型[199]、基于最小行为的恶意程序分析方法[200]、基于沙盒技术的恶意程序检测模型[201]等。

除研究恶意程序的检测技术外，也有很多学者研究恶意程序的传播问题。Peng等人[202]系统综述了智能手机恶意程序的分类和相应的传播模型。王长广等人[203]研究蓝牙环境下恶意程序的传播问题，将蓝牙协议的作用及设备移动方式抽象为不同的统计学参数，建立了一种蓝牙环境下恶意程序传播机制的分析模型。李婵婵等人[204]针对动态小世界社团网络病毒具有的社团结构和小世界特性，提出结构强度可调的动态小世界网络模型，来模拟现实生活中的本地接触和移动接触现象，并基于平均场理论建立了该网络上的SIR病毒传播模型。左黍和宋玉蓉[205]在考虑连边保护的情况下利用SIS模型建立了一种自适应网络病毒传播模型。林昭文等人[206]针对物联网中AS级网络拓扑结构，利用加权复杂网络提出了一种新的蠕虫传播模型。徐小龙等人[207]研究P2P网络中恶意代码的主动传播和被动传播两种情况，将P2P网络节点分为易感染、已暴露、已感染和已免疫4种状态，给出了以微分方程表示的处于各种状态的P2P网络节点数量随时间变化的演化公式。Ramachandran和Sikdar[208]利用仓室模型(Compartmental Model)分析采用Gnutella协议的P2P网络恶意程序传播问题，得到了控制恶意程序传播的系统参数和策略。Shan等人[209]利用强制存取控制(Mandatory Access Control)技术研究商业操作系统中的恶意程序防御问题，提出了一种能检测、跟踪且限制恶意入侵者的强制存取控制增强方法。Peng等人[210]提出了针对智能手机恶意程序的二维元胞自动机传播模型。Song等人[211]利用元胞自动机理论针对自适应网络恶意程序建立了传播模型。Yu等人[212]从网络全局考虑，将恶意程序传播分成两个阶段，再利用传染病理论分别建立不同阶段的恶意程序传播模型，揭示了恶意程序在不同阶段传播时分别具有指数分布和幂律分布的规律。Feng等人[213]在考虑

暂时免疫和传染率可变的情况下提出了一种时延 SIRS 模型，得到了决定恶意程序是否灭绝的重生数(Reproductive Number)。Khosroshahy 等人[214]在扩展传统传染病理论基础上提出了一种适用于 Botnet 的 SIC(Susceptible-Infected-Connected)模型，与其他恶意程序传播模型不同的是，该模型只跟踪已感染和已连接状态，并在考虑群体大小随机变化的基础上转化为连续时间马尔可夫链模型。Adu-Gyamfi 等人[215]针对移动社会网络恶意程序提出了 SEIRI(Susceptible-Exposed-Infected-Removed-Immune)模型。Wen 等人[216]提出了一种能维护时空同步过程的 SII(Susceptible-Infected-Immune)模型，该模型考虑了 Internet 用户检查邮件和社会信息的不一致性，以及节点与邻居节点之间的空间依赖性。Bose 和 Shin[217]在考虑网络用户应用程序交互、本地网络结构、用户移动性、恶意程序合作性等基础上，针对异质网络提出了一种基于智能体的恶意程序传播模型。Faghani 和 Nguyen[218]在考虑在线社会网络中的用户行为、社区的高聚簇结构、社区大小等因素基础上，针对 XSS(Cross-Site Scripting)恶意程序给出了节点被感染的概率模型。Wang 等人[219]针对车联 Ad Hoc 网络(Vehicular Ad Hoc Networks)，给出了一种基于城市均衡车流量，能反映移动特性、通信信道、介质访问控制机制的蠕虫传播模型。Karyotis 和 Papavassiliou[220]利用排队论建立了针对节点动态变化的复杂网络中恶意程序传播的模型。Cheng 等人[221]在考虑智能手机恶意程序具有的离域感染(Delocalized Infection)和波浪式传播(Ripple-based Propagation)特性基础上，给出了一种描述传播动态的微分方程模型。Lu 等人[222]提出了基于随机过程的 Botnet 恶意程序传播模型，得到了传播过程中节点移动性起决定性作用的结论。Peng 等人针对智能手机恶意程序，分别提出了基于半马尔可夫过程和社会关系图[223]、最具影响节点[224]的传播模型。

当前，越来越多的研究者开始关注无线传感器网络中具有自我复制功能的恶意程序的传播问题。付帅等人[225]考虑传感器节点的休眠与唤醒机制，提出一种 SIR/WS 模型，描述了感染传感器节点以广播方式传播恶意程序的过程，发现降低传染率和提高免疫率都可抑制恶意程序的传播。王小明等人[226]依据移动无线传感器网络的信息扩散机制，设计节点移动模型、无线信道分配算法、信号干扰模型和恶意数据包传播模型，定义移动传感器网络环境下的元胞空间、元胞邻域、元胞状态以及状态转换规则，提出移动无线传感器网络中恶意数据包传播的随机元胞自动机模型，研究了在不确定环境下恶意数据包传播的时空特征。他们还分别利用“反应扩散方程”(Reaction-diffusion Equation)[227]、脉冲微分方程(Pulse Differential Equation)[228]建立了恶意程序传播模型。Giannetsos 等人[229]说明了恶意程序如何在基于冯·诺依曼(Von Neumann)结构的传感器节点上传播的过程。通过将恶意代码分割成多个数据包，一个攻击者可以随意地将恶意程序注入传感器节点并完全控制这个传感器节点。然后注入的恶意程序就可以继续将自己分割成多个数据包并像自我复制的蠕虫一样以多跳方式传播自己直至整个无线传感器网络。为了抑制这种恶意程序的传播，研究者们提出了不同的模型来揭示恶意程序的传播机制。Khayam 和 Radha[230]利用信号处理技术建立了一种体现蠕虫时间和空间传播动力学的拓扑感知蠕虫传播模型(Topologically Aware Worm Propagation Model)，该模型同时考虑了物理通道条件、MAC 层碰撞、网络层路由和传输层协议的影响。Yanmaz[231]利用传统的 SIR 模型研究包含大量移动节点已被恶意程序感染后的静态无线网络，并根据网络的物理拓扑和移动节点模型，给出了恶意程序传播的规模和阈值。De 等人[232]提出一种流行病模型分析了 Trickle、

Firecracke、Deluge、MNP 等典型广播协议中的恶意程序传播率(Propagation Rate)等问题。利用相同的数学工具,他们[233]还分析了传感器节点欺骗的传播问题,并给出了其中的关键因素,这些因素将影响恶意程序传播的大爆发,从而避免可能导致的整个网络瘫痪。Mishra 和 Jha[234]提出了一种可能适用于描述无线传感器网络恶意程序传播的流行病模型 SEIQRS,该模型考虑了易感(Susceptible)、潜伏(Exposed)、感染(Infected)、隔离(Quarantined)、康复(Recovered)等不同状态。Wang 和 Li[235]在考虑传感器节点具有死亡状态(Dead State,即失去功能)的基础上提出了一种新的流行病传播模型 iSIRS 用于描述无线传感器网络中的恶意程序传播问题。Wang 等人[236]进一步提出另一种新的流行病传播模型 EiSIRS 以体现传感器节点具有休眠和工作交替调度的情况。另外,Tang[237]通过加入休眠状态改进了传统的 SI 模型,在该模型中引入了一种系统反病毒程序能在节点从工作状态到休眠状态转换的瞬间能自动启动并检查易感节点和查杀恶意程序的机制。

微分博弈作为一种研究多个具有利益纷争的理性参与者之间动态交互的博弈类型,在不同的无线网络环境中已有一些应用。Cao 等人[238]为了使无线传感器网络实现有效的监管目标,应用微分博弈得到了最优的追逃控制策略、信息需求的边界和这些边界的扩展属性。Miao 等人[239]针对能量限制的无线网络,利用合作微分博弈在权衡网络吞吐量(Network Throughput)和能量使用效率之间利益得失的基础上给出了一种优化的控制策略。与 Miao 等人[239]不同的是,Lin 等人[240]将非合作微分博弈应用于认知无线电 Ad Hoc 网络中多路径路由效率的提高上,得到的均衡经过证明是一种有效的路由分配方案。Zhu 等人[241]考虑异构无线网络中不同服务的带宽分配问题,为不同服务提供者之间的动态竞争关系建立了一种最大化带宽分配的微分博弈(Upper-bandwidth-allocation Differential Game)模型,并给出该模型的开环纳什均衡解。另外,Gu[242]利用微分博弈探索基于无线传感器网络的移动目标跟踪问题,通过计算相应博弈模型的鞍点均衡(Saddle-point Equilibrium),为估计要跟踪目标的位置提供了最优的筛选方案。Xu 和 Zhou[243]利用微分博弈研究低轨道卫星移动通信系统中的信道资源分配问题,给出了基于纳什均衡的卫星光束(Satellite Beam)优化分配方案。针对异构无线网络多路径路由中各条路由为追逐各自利益相互竞争有限的无线资源,导致数据传输的不可靠问题,Hu 和 Xie[244]利用非合作随机微分博弈(Noncooperative Stochastic Differential Game),以利益最大化为设计目标、有限带宽资源为限制条件、路径可靠度为关键因素,给出了基于反馈纳什均衡解的优化多路径路由策略。

然而,对于恶意程序传播过程中反映出来的决策问题,当前仅有少量研究利用博弈论方法解决该问题。Theodorakopoulos 等人[245]结合经典的流行病模型 SIR 和完全信息静态博弈用于解决动态部署网络的安全机制问题。Omic 等人[246]采用与 Theodorakopoulos 等人[245]相同的思路,但使用经典的 SIS 流行病模型解决如何在网络恶意程序传染时的网络保护问题。Bensoussan 等人[247]在僵尸网络(Botnet)环境中,利用一种改进的 SIS 模型研究被感染计算机(Bot)比例的动态演化问题,并通过微分博弈从经济利益的角度分析 Botnet 何时活动和系统如何采取最优防御策略的问题,最后在考虑防御策略的有效性和恶意程序传染率等因素基础上给出了两种可能的闭环纳什均衡解。另外,Khouzani 等人[248]将一种类似流行病模型(Epidemic-Similar Model)和微分博弈结合用于研究无线网络中的蠕虫传播问题,当相应的博弈模型达到鞍点后,得到一种鲁棒的防御策略,这种策略能根据被感染主

机比例的变化而动态地进行改变。

与上述相关工作相比，本章集中关注无线传感器网络中的恶意程序传播问题，而Bensoussan等人[247]和Khouzani等人[248]分别关注僵尸网络和通常的无线网络环境。本章与Wang等人[236]研究工作类似的是都研究无线传感器网络环境，但本章考虑了无线传感器网络系统与恶意程序的主观努力程度(Effort Intensity)，因此得到了不同的恶意程序传播模型，尤其在表达传感器节点状态动态演化的微分等式方面。本章将使用微分博弈建立恶意程序防御微分博弈模型，该模型的研究目标是在恶意程序动态改变其策略时，为无线传感器网络系统得到动态的最优防御策略。这种融合流行病理论和博弈论的思想类似于Bensoussan等人[247]和Khouzani等人[248]的工作。然而，本章为满足无线传感器网络的特点，扩展的流行病模型考虑每个传感器节点具有 S(易感)、$\widetilde{S}$(易感且休眠)、I(被感染)、$\widetilde{I}$(被感染且休眠)、R(康复)、$\widetilde{R}$(康复且休眠)、D(死亡)等7种状态，而Bensoussan等人[247]仅考虑了 S 和 I 两种状态，Khouzani等人[248]考虑了 S、I、R 和 D 等4种状态，因此本章得到的最优控制策略与他们[247, 248]得到的最优控制策略完全不同。

5.3 基于扩展流行病理论的无线传感器网络恶意程序传播模型

本章考虑的无线传感器网络环境包含 N 个静态已标识的传感器节点，这些传感器节点以节点密度 σ 被统一地撒播在一块二维的区域上。每个传感器节点都已配备具有最大信号发射距离为 r 的双向天线。从源传感器节点捕获的数据能被发送到信号传输范围内的相邻传感器节点，这些邻居传感器节点再以中继方式继续将数据传递到下一个相邻的传感器节点。另外，所有传感器节点的供电都采用有限电能的电池。

以流行病理论的观点来看，无线传感器网络中的传感器节点因其自身的特性，可以分成以下的7种状态：

① S(Susceptible)。处于状态 S 的传感器节点正在正常工作且易于被恶意程序感染，但当前还未被恶意程序感染。

② $\widetilde{S}$(Susceptible and Sleeping)。处于状态 $\widetilde{S}$ 的传感器节点虽然存在被恶意程序利用的漏洞，但因正在休眠而无通信功能，所以，它不会被恶意程序感染。

③ I(Infected)。处于状态 I 的传感器节点已被恶意程序感染且能通过传输数据或控制信息的方式将恶意程序传染到处于状态 S 的相邻传感器节点上。

④ $\widetilde{I}$(Infected and Sleeping)。处于状态 $\widetilde{I}$ 的传感器节点虽然已被恶意程序感染，但因正在休眠而无通信功能，所以它不能将恶意程序传染到其他相邻的传感器节点上。

⑤ R(Recovered)。处于状态 R 的传感器节点可能是因为安装了安全补丁后从状态 S 中转换过来，也可能经过治疗后从状态 I 中转换过来，这种传感器节点对恶意程序具有免疫力，也就是说不会被恶意程序感染。

⑥ $\widetilde{R}$(Recovered and Sleeping)。处于状态 $\widetilde{R}$ 的传感器节点对恶意程序有免疫力且正在休眠。

⑦ D(Dead)。处于状态 D 的传感器节点已失去所有的功能，它们可能因为能量消耗殆

尽而从状态 S、I 和 R 转换过来，也可能因为恶意程序故意破坏而从状态 I 转换过来。当然，这样的传感器节点即使已被恶意程序感染，也不会传播恶意程序。

图 5-1 给出了传感器节点的所有 7 种状态之间的转换关系。

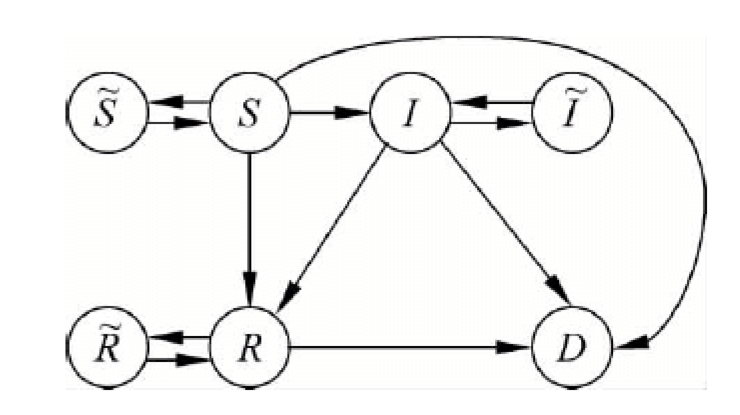


图 5-1　传感器节点状态间的动态演化

在图 5-1 中，假设

$$\forall\, t, S(t)+\widetilde{S}(t)+I(t)+\widetilde{I}(t)+R(t)+\widetilde{R}(t)+D(t)=N \tag{5-1}$$

始终成立，其中 $\cdot(t)$ 表示在时刻 t 处于状态 $\cdot$ 的传感器节点数量。另外，还假设在恶意程序大规模传染爆发前已有部分传感器节点已被恶意程序感染，即

$$0<I(0)=I_0<N \tag{5-2}$$

式中，I_0 为初始已感染恶意程序节点数。为简化起见，令

$$\widetilde{S}(0)=\widetilde{I}(0)=R(0)=\widetilde{R}(0)=D(0)=0 \tag{5-3}$$

从而

$$S(0)=N-I_0 \tag{5-4}$$

接下来本章将以流行病理论的观点分析各种状态间的动态变化关系。

对处于状态 S 的传感器节点而言，当它与已被恶意程序感染的传感器节点通信时，它被恶意程序感染的概率为 γ_{SI}。由于所有的传感器节点以节点密度 σ 均匀地被撒播在传感区域中，因此一个被恶意程序感染的传感器节点能与它通信的相邻传感器节点数为 $\sigma\pi r^2$。然而，这些所有相邻的传感器节点不是都会被恶意程序感染，只有那些处于状态 S 的传感器节点才有可能被感染。由前文所述传感器节点在传感区域内统一分布的假设，可得在时刻 t 相邻的传感器节点是易感传感器节点的概率为 $S(t)/N$。因此，所有的已感染传感器节点和易感传感器节点形成的节点对数目为 $\sigma\pi r^2 S(t)I(t)/N$。这样，就可得到从状态 S 转换到状态 I 的传感器节点数为 $\gamma_{SI}\sigma\pi r^2 S(t)I(t)/N$。另外，为了节省传感器节点的能量，传感器节点以概率 $\gamma_{S\widetilde{S}}$ 进入休眠状态。因此，在时刻 t 从状态 S 转换到状态 $\widetilde{S}$ 的传感器节点数为 $\gamma_{S\widetilde{S}}S(t)$。类似地，可以得到从状态 S 转换到状态 R 和 D 的传感器节点数分别为 $\gamma_{SR}S(t)$ 和 $\gamma_{SD}S(t)$，其中 γ_{SR} 表示一个处于状态 S 的传感器节点被成功安装安全补丁以便对恶意程序具有免疫力的概率，γ_{SD} 表示一个状态 S 的传感器节点因能量消耗殆尽而失去所有功能的概率。

对处于状态 I 的传感器节点而言，无线传感器网络系统会以概率 $\gamma_{I\widetilde{I}}$ 控制其进入休眠状态以避免其将恶意程序传染给其他传感器节点。由于安全补丁能清除处于状态 I 的传感器节点上的恶意程序，因此，这种被感染的传感器节点具有转换为康复传感器节点的概率，记为 γ_{IR}。因为从状态 $\widetilde{S}$ 和 I 转换到状态 R 都与安装安全补丁有关，为简便起见，假设 $\gamma_{IR}=\gamma_{SR}$。另外，这种被感染传感器节点可能会因为能量消耗殆尽而失去所有功能，也有可能被

恶意程序故意杀死，因此从状态 I 转换到状态 D 具有一定的概率，记为 γ_{ID}。所以，可得到在时刻 t 被感染传感器节点转换为状态$\widetilde{I}$、R 和 D 的数量分别为 $\gamma_{I\widetilde{I}}I(t)$、$\gamma_{IR}I(t)$和 $\gamma_{ID}I(t)$。

对处于状态 R 的传感器节点而言，由于正常的系统调度进入休眠状态，所以转换到状态$\widetilde{R}$具有一定的概率，记为 $\gamma_{R\widetilde{R}}$；又由于其能量会耗尽而失去所有功能，所以转换到状态 D 具有一定的概率，记为 γ_{RD}，为简化起见，假设 $\gamma_{RD}=\gamma_{SD}$。所以，可得到在时刻 t 处于状态 R 的传感器节点转换到状态$\widetilde{R}$ 和 D 的数量分别为 $\gamma_{R\widetilde{R}}R(t)$和 $\gamma_{RD}R(t)$。

对处于状态$\widetilde{S}$、$\widetilde{I}$ 和$\widetilde{R}$中的传感器节点而言，它们会在系统的正常控制下被唤醒，这些转换具有一定的概率，分别记为 $\gamma_{\widetilde{S}S}$、$\gamma_{\widetilde{I}I}$和 $\gamma_{\widetilde{R}R}$。为简化起见，假设 $\gamma_{\widetilde{S}S}=\gamma_{\widetilde{I}I}=\gamma_{\widetilde{R}R}$。这样，可得到在时刻 t 处于状态$\widetilde{S}$、$\widetilde{I}$ 和$\widetilde{R}$中的传感器节点转换到状态 S、I 和 R 的数量分别为 $\gamma_{\widetilde{S}S}\widetilde{S}(t)$、$\gamma_{\widetilde{I}I}\widetilde{I}(t)$和 $\gamma_{\widetilde{R}R}\widetilde{R}(t)$。

实际上，在上面引入的概率参数中，γ_{SR}、γ_{IR}和 $\gamma_{I\widetilde{I}}$受无线传感器网络系统控制，而 γ_{SI}和 γ_{ID}由恶意程序控制，这些参数在整个无线传感器网络部署完毕后，属于一些静态的数值，因此不能反映出无线传感器网络系统和恶意程序各自的主观努力程度，为此，本章为无线传感器网络系统引入在时刻 t 相应的主观努力度参数 $\gamma_{SR}^{S}(t)$、$\gamma_{IR}^{S}(t)$和 $\gamma_{I\widetilde{I}}^{S}(t)$，为恶意程序引入在时刻 t 相应的主观努力度参数 $\gamma_{SI}^{M}(t)$和 $\gamma_{ID}^{M}(t)$。其中 $\gamma_{SR}^{S}(t)$、$\gamma_{IR}^{S}(t)$和 $\gamma_{I\widetilde{I}}^{S}(t)$由无线传感器网络系统根据恶意程序传播的状况动态控制，而 $\gamma_{SI}^{M}(t)$和 $\gamma_{ID}^{M}(t)$由恶意程序根据无线传感器网络系统采取的策略进行相应的动态控制。由于 $\gamma_{SR}^{S}(t)$和 $\gamma_{IR}^{S}(t)$都与安全补丁的安装有关，为简化起见，假设 $\gamma_{SR}^{S}(t)=\gamma_{IR}^{S}(t)$。在考虑这些引入的主观努力度参数后，就可以分别重写在时刻 t 从状态 S 转换到状态 I 的数量为 $\gamma_{SI}^{M}(t)\gamma_{SI}\sigma\pi r^2S(t)I(t)/N$；从状态 S 转换到状态 R 的数量为 $\gamma_{SR}^{S}(t)\gamma_{SR}S(t)$；从状态 I 转换到状态 R 的数量为 $\gamma_{IR}^{S}(t)\gamma_{IR}I(t)$；从状态 I 转换到状态$\widetilde{I}$ 的数量为 $\gamma_{I\widetilde{I}}^{S}(t)\gamma_{I\widetilde{I}}I(t)$；从状态 I 转换到 D 的数量为 $\gamma_{ID}^{M}(t)\gamma_{ID}I(t)$。除了这些主观努力度参数外，其他的参数由于不是防御恶意程序传播的关键因素且只与无线传感器网络的静态部署有关，因此可以将它们考虑成常数。

经过上述分析，可得到无线传感器网络传感器节点各状态动态变化的微分等式。对

$$\forall t,\ S(t), \widetilde{S}(t), I(t), \widetilde{I}(t), R(t), \widetilde{R}(t), D(t) \geqslant 0 \tag{5-5}$$

且

$$S(t)+\widetilde{S}(t)+I(t)+\widetilde{I}(t)+R(t)+\widetilde{R}(t)+D(t)=N \tag{5-6}$$

有

$$\begin{cases} \dfrac{\mathrm{d}S(t)}{\mathrm{d}t}=-\gamma_{SI}^{M}(t)\gamma_{SI}\sigma\pi r^2S(t)I(t)/N-\gamma_{S\widetilde{S}}S(t)+\gamma_{\widetilde{S}S}\widetilde{S}(t)-\gamma_{SR}^{S}(t)\gamma_{SR}S(t)-\gamma_{SD}S(t) \\ S(0)=N-I_0 \end{cases} \tag{5-7}$$

$$\begin{cases} \dfrac{\mathrm{d}\widetilde{S}(t)}{\mathrm{d}t}=\gamma_{S\widetilde{S}}S(t)-\gamma_{\widetilde{S}S}\widetilde{S}(t) \\ \widetilde{S}(0)=0 \end{cases} \tag{5-8}$$

$$\begin{cases} \dfrac{\mathrm{d}I(t)}{\mathrm{d}t} = \gamma_{SI}^{M}(t)\gamma_{SI}\sigma\pi r^2 S(t)I(t)/N - \gamma_{I\tilde{I}}^{S}(t)\gamma_{I\tilde{I}}I(t) + \gamma_{\tilde{I}I}\tilde{I}(t) - \gamma_{IR}^{S}(t)\gamma_{IR}I(t) - \gamma_{ID}^{M}(t)\gamma_{ID}I(t) \\ I(0) = I_0 \end{cases} \tag{5-9}$$

$$\begin{cases} \dfrac{\mathrm{d}\tilde{I}(t)}{\mathrm{d}t} = \gamma_{I\tilde{I}}^{S}(t)\gamma_{I\tilde{I}}I(t) - \gamma_{\tilde{I}I}\tilde{I}(t) \\ \tilde{I}(0) = 0 \end{cases} \tag{5-10}$$

$$\begin{cases} \dfrac{\mathrm{d}R(t)}{\mathrm{d}t} = \gamma_{SR}^{S}(t)\gamma_{SR}S(t) + \gamma_{IR}^{S}(t)\gamma_{IR}I(t) - \gamma_{R\tilde{R}}R(t) + \gamma_{\tilde{R}R}\tilde{R}(t) - \gamma_{RD}R(t) \\ R(0) = 0 \end{cases} \tag{5-11}$$

$$\begin{cases} \dfrac{\mathrm{d}\tilde{R}(t)}{\mathrm{d}t} = \gamma_{R\tilde{R}}R(t) - \gamma_{\tilde{R}R}\tilde{R}(t) \\ \tilde{R}(0) = 0 \end{cases} \tag{5-12}$$

$$\begin{cases} \dfrac{\mathrm{d}D(t)}{\mathrm{d}t} = \gamma_{ID}^{M}(t)\gamma_{ID}I(t) + \gamma_{RD}R(t) + \gamma_{SD}S(t) \\ D(0) = 0 \end{cases} \tag{5-13}$$

5.4 基于微分博弈的最优控制策略

5.4.1 无线传感器网络恶意程序防御微分博弈模型

定义 5-1 给定一固定逗留期 T，无线传感器网络“恶意程序防御微分博弈”是零和的且由一个四元组$\mathbb{G}=(\mathcal{N},\mathcal{C},\mathcal{F},J)$组成，其中：

- $\mathcal{N}$={无线传感器网络系统，恶意程序}是博弈的参与者集合。
- $\mathcal{C}=\mathcal{U}\times\mathcal{V}$，其中，

  $\mathcal{U}=\{\mu(t)\,|\,\mu(t)=(\gamma_{IR}^{S}(t),\ \gamma_{I\tilde{I}}^{S}(t)),\ 0\leqslant\gamma_{IR}^{S}(t)\leqslant 1,\ 0\leqslant\gamma_{I\tilde{I}}^{S}(t)\leqslant 1\}$是参与者无线传感器网络系统在时刻 t 可用的控制策略集合。

  $\mathcal{V}=\{\nu(t)\,|\,\nu(t)=(\gamma_{SI}^{M}(t),\ \gamma_{ID}^{M}(t)),\ 0\leqslant\gamma_{SI}^{M}(t)\leqslant 1,\ 0\leqslant\gamma_{ID}^{M}(t)\leqslant 1\}$是参与者恶意程序在时刻 t 可用的控制策略集合。
- $\mathcal{F}=\{f(t,\boldsymbol{x}(t),\mu(t),\nu(t))\,|\,f_S,f_{\tilde{S}},f_I,f_{\tilde{I}},f_R,f_{\tilde{R}},f_D\}$，其中 $f(t,\boldsymbol{x}(t),\mu(t),\nu(t))$是一个状态函数且 $\boldsymbol{x}(t)=[S(t)\ \ \tilde{S}(t)\ \ I(t)\ \ \tilde{I}(t)\ \ R(t)\ \ \tilde{R}(t)\ \ D(t)]$是一个七维的状态向量。所有的状态函数分别由式(5-7)至式(5-13)确定，即 $f_S=\dfrac{\mathrm{d}S(t)}{\mathrm{d}t}$、$f_{\tilde{S}}=\dfrac{\mathrm{d}\tilde{S}(t)}{\mathrm{d}t}$、$f_I=\dfrac{\mathrm{d}I(t)}{\mathrm{d}t}$、$f_{\tilde{I}}=\dfrac{\mathrm{d}\tilde{I}(t)}{\mathrm{d}t}$、$f_R=\dfrac{\mathrm{d}R(t)}{\mathrm{d}t}$、$f_{\tilde{R}}=\dfrac{\mathrm{d}\tilde{R}(t)}{\mathrm{d}t}$、$f_D=\dfrac{\mathrm{d}D(t)}{\mathrm{d}t}$。
- $$J(\mu(t),\nu(t)) = \int_0^T g(t,\boldsymbol{x}(t),\mu(t),\nu(t))\mathrm{d}t + q(\boldsymbol{x}(T)) \tag{5-14}$$

这是一个反映无线传感器网络 QoS 的成本函数，其中 $g:[0,T]\times\mathbb{R}^7\times\mathcal{U}\times\mathcal{V}\mapsto\mathbb{R}$ 代表在时刻 t 的瞬时成本(Instantaneous Cost)，而 $q:\boldsymbol{x}(t)\mapsto\mathbb{R}$ 表示逗留期结束时的终期成本

(Terminal Cost)。

在定义 5-1 中,考虑整个微分博弈中的参与者由无线传感器网络系统和恶意程序组成。其中"无线传感器网络系统"实质是管理传感器节点和控制它们行为的系统软件。在一个无线传感器网络中可能存在多种类型的恶意程序,这些恶意程序会窃听由传感器节点感知的秘密的数据信息,甚至破坏整个无线传感器网络的正常通信,直至使整个网络瘫痪。由于它们具有共同点,用参与者恶意程序可以代表所有的恶意程序。定义无线传感器网络恶意程序防御微分博弈的目的就是要从博弈论的角度在参与者恶意程序动态改变传播策略时为参与者无线传感器网络系统提供最优的控制策略。

本章分别选择($\gamma_{IR}^{S}(t)$,$\gamma_{I\tilde{I}}^{S}(t)$)和($\gamma_{SI}^{M}(t)$, $\gamma_{ID}^{M}(t)$)作为参与者无线传感器网络系统和恶意程序的控制(Control)参数,这种选择是由无线传感器网络恶意程序传播的影响因素来确定的。参数 $\gamma_{IR}^{S}(t)$反映了无线传感器网络系统分发安全补丁的主观努力程度,其值越大,传感器节点通过安装安全补丁从状态 I 转换为状态 R 的数量越大,这样使得能对恶意程序免疫的传感器节点越多,进而增强整个无线传感器网络通信的稳定性。同时,较高的 $\gamma_{I\tilde{I}}^{S}(t)$值将使更多的被感染传感器节点进入休眠状态,从而使这些传感器节点失去传播恶意程序的能力。另外,参数 $\gamma_{SI}^{M}(t)$代表了恶意程序将传感器节点从状态 S 转换为状态 I 的主观努力程度,其值越高,传播恶意程序的机会越多。恶意程序通过执行特殊的代码可以使被感染传感器节点丧失所有功能,这主要通过参数 $\gamma_{ID}^{M}(t)$来反映其主观努力程度。这两个因素都将干扰传感器节点的正常通信,从而降低无线传感器网络的 QoS。

接下来分析与参与者无线传感器网络系统和恶意程序相关的成本函数。在被恶意程序感染的传感器节点上,恶意程序可以干扰传感器节点的正常工作甚至毁坏传感器节点。同时,恶意程序能收集传感器节点感知的数据从而引起隐私信息泄露,也能对传感器系统资源进行非授权的访问。因此,这些感染恶意程序的传感器节点在时刻 t 会引起一个瞬时成本 $c_I I(t)$,其中 c_I 为系数且 $c_I \geqslant 0$。这里要说明的是,瞬时成本采用线性表达式进行描述是为了方便后期的计算。虽然使用非线性表达式进行描述可能会更精确,但任何非线性表达式都可以近似地使用线性表达式表示出来,因此这种表达瞬时成本的方式是合理的。对那些被恶意的程序故意杀死的传感器节点,它们将失去所有功能且中断所有通过它们的路由,这种恶意程序的行为将在时刻 t 对整个无线传感器网络产生瞬时成本 $c_D D(t)$,其中 c_D 为一个系数且 $c_D \geqslant 0$。更进一步,这些被恶意程序杀死的传感器节点还会产生一个终期成本 $c_D^{T} D(t)$,其中 c_D^{T} 为一个系数且 $c_D^{T} \geqslant 0$,这主要因为那些被杀死的传感器节点需要购买新的传感器节点从而产生成本。对那些经过安装安全补丁的康复传感器节点,它们已具有对恶意程序的免疫力,这种状况将有益于整个无线传感器网络,因此需要减去在时刻 t 产生的瞬时成本 $c_{IR}\gamma_{IR}^{S}(t)\gamma_{IR}$,其中 c_{IR} 为一个系数且 $c_{IR} \geqslant 0$。这里仅选择 $\gamma_{IR}^{S}(t)$作为瞬时成本的一个变量,实际上已包括了参数 $\gamma_{SR}^{S}(t)$的影响,因为这两个参数 $\gamma_{SR}^{S}(t)$和 $\gamma_{IR}^{S}(t)$都与安全补丁的安装有关且已假设 $\gamma_{SR}^{S}(t)=\gamma_{IR}^{S}(t)$。然而,安装安全补丁的过程将扫描所有的传感器节点并传输相应的安全补丁,因此会在时刻 t 产生一个瞬时成本 $c_R R(t)$,其中 c_R 为一个系数且 $c_R \geqslant 0$。对那些已感染恶意程序并被无线传感器网络系统控制进入休眠状态的传感器节点,它们失去原有传播恶意程序的能力,因此整个无线传感器网络会从这些转换中获得收益,也就是说,要减去在时刻 t 产生的瞬时成本 $c_{I\tilde{I}}\gamma_{I\tilde{I}}^{S}(t)\gamma_{I\tilde{I}}$,其中 $c_{I\tilde{I}}$是一个系数且 $c_{I\tilde{I}} \geqslant 0$。过多的传感器节点进入休眠状态将会使传感器网络的通信受阻,因此将在时刻 t 产生相应的瞬时成本 $c_{\tilde{S}}\tilde{S}(t)$

和 $c_{\widetilde{R}}\widetilde{R}(t)$，其中 $c_{\widetilde{S}}\geqslant 0$、$c_{\widetilde{R}}\geqslant 0$。而隔离被感染的传感器节点（即控制被感染传感器节点进入休眠状态）失去了传播恶意程序的能力，因此需要减去在时刻 t 产生的瞬时成本 $c_{\widetilde{I}}\widetilde{I}(t)$，其中 $c_{\widetilde{I}}$ 是一个系数且 $c_{\widetilde{I}}\geqslant 0$。根据上述分析，可重写式(5-14)为

$$J(\mu(t),\nu(t))=\int_0^T(c_I I(t)+c_D D(t)+c_R R(t)+c_{\widetilde{S}}\widetilde{S}(t)-c_{\widetilde{I}}\widetilde{I}(t)+c_{\widetilde{R}}\widetilde{R}(t)-c_{IR}\gamma_{IR}^{S}(t)\gamma_{IR}-c_{I\widetilde{I}}\gamma_{I\widetilde{I}}^{S}(t)\gamma_{I\widetilde{I}})\,\mathrm{d}t+c_D^{T}D(T) \tag{5-15}$$

得到无线传感器网络 QoS 的成本函数后，无线传感器网络系统面临的问题是如何选择它的最优控制策略 $\mu^*(t)$ 来最小化式(5-15)，而恶意程序是如何选择它的最优控制策略 $\nu^*(t)$ 来最大化式(5-15)。解决这些问题可从寻找无线传感器网络恶意程序防御微分博弈的鞍点入手。

定理 5-1 无线传感器网络恶意程序防御微分博弈$\mathbb{G}$存在鞍点。

证明 根据定义 5-1，可得到：

(1) 状态函数 $f(t,\boldsymbol{x}(t),\mu(t),\nu(t))$ 在状态空间 $S(t)$、$\widetilde{S}(t)$、$I(t)$、$\widetilde{I}(t)$、$R(t)$、$\widetilde{R}(t)$ 和 $D(t)$ 及控制空间 $\gamma_{SR}^{S}(t)$、$\gamma_{IR}^{S}(t)$、$\gamma_{I\widetilde{I}}^{S}(t)$、$\gamma_{SI}^{M}(t)$ 和 $\gamma_{ID}^{M}(t)$ 上连续且有界。

(2) 瞬时成本函数 $g(t,\boldsymbol{x}(t),\mu(t),\nu(t))$ 和终期成本函数 $q(\boldsymbol{x}(T))$ 在所有的状态和控制空间上是连续的。

(3) 状态函数 $f(t,\boldsymbol{x}(t),\mu(t),\nu(t))$ 和瞬时成本函数 $g(t,\boldsymbol{x}(t),\mu(t),\nu(t))$ 是以控制参数线性表达的。

(4) 参与者无线传感器网络系统的控制策略集合与参与者恶意程序控制策略集合形成的笛卡儿积集合$\mathcal{C}$是凸的。

上述的 4 个条件满足文献[249]中定理 2.6 需要满足的条件，因此定理 5-1 成立。证毕。

定理 5-1 意味着无线传感器网络恶意程序防御微分博弈$\mathbb{G}$存在一对鞍点控制策略 $(\mu^*(t),\nu^*(t))$ 满足

$$J(\mu^*(t),\nu(t))\leqslant J(\mu^*(t),\nu^*(t))\leqslant J(\mu(t),\nu^*(t)) \tag{5-16}$$

这就是说，当参与者无线传感器网络系统选择鞍点策略 $\mu^*(t)$ 时，不管参与者恶意程序选择何种控制策略，对无线传感器网络系统而言，至多产生成本 $J(\mu^*(t),\nu^*(t))$；而当参与者恶意程序选择鞍点策略 $\nu^*(t)$ 时，不管参与者无线传感器网络系统选择何种控制策略，对恶意程序而言至少产生 $J(\mu^*(t),\nu^*(t))$。定理 5-1 的结论满足文献[36]中定理 2.7 的条件，因此无线传感器网络恶意程序防御微分博弈$\mathbb{G}$存在值 V 使得

$$V=V^-=V^+=J(\mu^*(t),\nu^*(t)) \tag{5-17}$$

其中，

$$V^-=\max_{\nu(t)}\min_{\mu(t)}J(\mu(t),\nu(t)) \tag{5-18}$$

$$V^+=\min_{\mu(t)}\max_{\nu(t)}J(\mu(t),\nu(t)) \tag{5-19}$$

因此，这种最小化最大可产生成本的鞍点控制策略 $\mu^*(t)$ 是参与者无线传感器网络系统的最优控制策略，而最大化最小可产生成本的鞍点控制 $\nu^*(t)$ 是参与者恶意程序的最优控制策略。

5.4.2 无线传感器网络系统和恶意程序的最优控制

设 $\boldsymbol{x}^*(t)$、$S^*(t)$、$I^*(t)$分别为无线传感器网络恶意程序防御微分博弈达到鞍点控制策略$(\mu^*(t),\nu^*(t))$，即无线传感器网络系统和恶意程序都分别达到最优控制 $\gamma_{IR}^{S*}(t)$、$\gamma_{I\widetilde{I}}^{S*}(t)$、$\gamma_{SI}^{M*}(t)$和 $\gamma_{ID}^{M*}(t)$时最优的状态轨迹、易感传感器节点的数量和被感染传感器节点的数量。由定理 5-1，无线传感器网络恶意程序防御微分博弈存在鞍点，因此存在一个协状态函数向量

$$\boldsymbol{p}(t)=[p_S(t)\ p_{\widetilde{S}}(t)\ p_I(t)\ p_{\widetilde{I}}(t)\ p_R(t)\ p_{\widetilde{R}}(t)\ p_D(t)]^{\mathrm{T}} \tag{5-20}$$

使得以下的必要条件成立，即

$$\begin{cases}\forall f\in\{f_S,f_{\widetilde{S}},f_I,f_{\widetilde{I}},f_R,f_{\widetilde{R}},f_D\}\\ \dfrac{\mathrm{d}\boldsymbol{x}^*(t)}{\mathrm{d}t}=f(t,\boldsymbol{x}^*(t),\mu^*(t),\nu^*(t)),\ \boldsymbol{x}^*(0)=\boldsymbol{x}_0\end{cases} \tag{5-21}$$

$$\begin{cases}\forall\mu(t)\in\mathcal{U},\forall\nu(t)\in\mathcal{V},\forall t\in[0,T]\\ H(t,\boldsymbol{p}(t),\boldsymbol{x}^*(t),\mu^*(t),\nu(t))\leqslant H(t,\boldsymbol{p}(t),\boldsymbol{x}^*(t),\mu^*(t),\nu^*(t))\\ \qquad\leqslant H(t,\boldsymbol{p}(t),\boldsymbol{x}^*(t),\mu(t),\nu^*(t))\end{cases} \tag{5-22}$$

$$\begin{cases}\dfrac{\mathrm{d}p_z(t)}{\mathrm{d}z}=-\dfrac{\partial}{\partial z}H(t,\boldsymbol{p}(t),\boldsymbol{x}^*(t),\mu^*(t),\nu^*(t))\\ p_z(T)=\dfrac{\partial}{\partial z}q(\boldsymbol{x}^*(T))\end{cases} \tag{5-23}$$

其中，哈密尔顿等式

$$\begin{aligned}&H(t,\boldsymbol{p}(t),\boldsymbol{x}(t),\mu(t),\nu(t))=g(t,\boldsymbol{x}(t),\mu(t),\nu(t))\\&\quad+\sum_{i\in\{S,\widetilde{S},I,\widetilde{I},R,\widetilde{R},D\}}p_i(t)f(t,\boldsymbol{x}(t),\mu(t),\nu(t))\\&=c_II(t)+c_DD(t)+c_RR(t)+c_{\widetilde{S}}\widetilde{S}(t)-c_{\widetilde{I}}\widetilde{I}(t)+c_{\widetilde{R}}\widetilde{R}(t)\\&\quad+p_S(t)(-\gamma_{S\widetilde{S}}S(t)+\gamma_{\widetilde{S}S}\widetilde{S}(t)-\gamma_{SD}S(t))\\&\quad+p_{\widetilde{S}}(t)(\gamma_{S\widetilde{S}}S(t)-\gamma_{\widetilde{S}S}\widetilde{S}(t))+p_I(t)\gamma_{\widetilde{I}I}\widetilde{I}(t)\\&\quad-p_{\widetilde{I}}(t)\gamma_{\widetilde{I}I}\widetilde{I}(t)+p_R(t)(-\gamma_{R\widetilde{R}}R(t)+\gamma_{\widetilde{R}R}\widetilde{R}(t)-\gamma_{RD}R(t))\\&\quad+p_{\widetilde{R}}(t)(\gamma_{R\widetilde{R}}R(t)-\gamma_{\widetilde{R}R}\widetilde{R}(t))+p_D(t)(\gamma_{RD}R(t)+\gamma_{SD}S(t))\\&\quad+\gamma_{IR}^S(t)\gamma_{IR}(p_R(t)(S(t)+I(t))-c_{IR}-p_S(t)S(t)-p_I(t)I(t))\\&\quad+\gamma_{I\widetilde{I}}^S(t)\gamma_{I\widetilde{I}}(-c_{I\widetilde{I}}-p_I(t)I(t)+p_{\widetilde{I}}(t)I(t))\\&\quad+\gamma_{SI}^M(t)\gamma_{SI}\sigma\pi r^2S(t)I(t)/N(p_I(t)-p_S(t))\\&\quad+\gamma_{ID}^M(t)\gamma_{ID}I(t)(p_D(t)-p_I(t))\\&\triangleq\varphi(t)+\gamma_{IR}^S(t)\eta_{IR}(t)+\gamma_{I\widetilde{I}}^S(t)\eta_{I\widetilde{I}}(t)+\gamma_{SI}^M(t)\eta_{SI}(t)+\gamma_{ID}^M(t)\eta_{ID}(t)\end{aligned} \tag{5-24}$$

需要注意的是，式(5-24)中的 $\gamma_{SR}^S(t)$和 γ_{SR}已分别用 $\gamma_{IR}^S(t)$和 γ_{IR}替换。根据式(5-15)、式(5-23)和式(5-24)，可得

$$\begin{cases}\dfrac{\mathrm{d}p_S(t)}{\mathrm{d}t}=-\dfrac{\partial}{\partial S(t)}H(t,\boldsymbol{p}(t),\boldsymbol{x}^*(t),\mu^*(t),\nu^*(t))\\ \qquad = p_S(t)(\gamma_{SI}^{M*}(t)\gamma_{SI}\sigma\pi r^2 I^*(t)/N+\gamma_{IR}^{S*}(t)\gamma_{IR}+\gamma_{S\tilde{S}}+\gamma_{SD})-p_{\tilde{S}}(t)\gamma_{S\tilde{S}}\\ \qquad\quad -p_I(t)\gamma_{SI}^{M*}(t)\gamma_{SI}\sigma\pi r^2 I^*(t)/N-p_R(t)\gamma_{IR}^{S*}(t)\gamma_{IR}-p_D(t)\gamma_{SD}\\ p_S(T)=0\end{cases}\tag{5-25}$$

$$\begin{cases}\dfrac{\mathrm{d}p_{\tilde{S}}(t)}{\mathrm{d}t}=-\dfrac{\partial}{\partial \tilde{S}(t)}H(t,\boldsymbol{p}(t),\boldsymbol{x}^*(t),\mu^*(t),\nu^*(t))\\ \qquad =-p_S(t)\gamma_{\tilde{S}S}+p_{\tilde{S}}(t)\gamma_{\tilde{S}S}-c_{\tilde{S}}\\ p_{\tilde{S}}(T)=0\end{cases}\tag{5-26}$$

$$\begin{cases}\dfrac{\mathrm{d}p_I(t)}{\mathrm{d}t}=-\dfrac{\partial}{\partial I(t)}H(t,\boldsymbol{p}(t),\boldsymbol{x}^*(t),\mu^*(t),\nu^*(t))\\ \qquad =-c_I+p_S(t)\gamma_{SI}^{M*}(t)\gamma_{SI}\sigma\pi r^2 S^*(t)/N\\ \qquad\quad -p_I(t)(\gamma_{SI}^{M*}(t)\gamma_{SI}\sigma\pi r^2 S^*(t)/N+\gamma_{I\tilde{I}}^{S*}(t)\gamma_{I\tilde{I}}+\gamma_{IR}^{S*}(t)\gamma_{IR}\\ \qquad\quad +\gamma_{ID}^{M*}(t)\gamma_{ID})-p_{\tilde{I}}(t)\gamma_{I\tilde{I}}^{S*}(t)\gamma_{I\tilde{I}}-p_R(t)\gamma_{IR}^{S*}(t)\gamma_{IR}-p_D(t)\gamma_{ID}^{M*}(t)\gamma_{ID}\\ p_I(T)=0\end{cases}\tag{5-27}$$

$$\begin{cases}\dfrac{\mathrm{d}p_{\tilde{I}}(t)}{\mathrm{d}t}=-\dfrac{\partial}{\partial \tilde{I}(t)}H(t,\boldsymbol{p}(t),\boldsymbol{x}^*(t),\mu^*(t),\nu^*(t))\\ \qquad =-p_I(t)\gamma_{\tilde{I}I}+p_{\tilde{I}}(t)\gamma_{\tilde{I}I}+c_{\tilde{I}}\\ p_{\tilde{I}}(T)=0\end{cases}\tag{5-28}$$

$$\begin{cases}\dfrac{\mathrm{d}p_R(t)}{\mathrm{d}t}=-\dfrac{\partial}{\partial R(t)}H(t,\boldsymbol{p}(t),\boldsymbol{x}^*(t),\mu^*(t),\nu^*(t))\\ \qquad = p_R(t)(\gamma_{R\tilde{R}}+\gamma_{RD})-p_{\tilde{R}}(t)\gamma_{R\tilde{R}}-p_D(t)\gamma_{RD}-c_R\\ p_R(T)=0\end{cases}\tag{5-29}$$

$$\begin{cases}\dfrac{\mathrm{d}p_{\tilde{R}}(t)}{\mathrm{d}t}=-\dfrac{\partial}{\partial \tilde{R}(t)}H(t,\boldsymbol{p}(t),\boldsymbol{x}^*(t),\mu^*(t),\nu^*(t))\\ \qquad =-p_R(t)\gamma_{\tilde{R}R}+p_{\tilde{R}}(t)\gamma_{\tilde{R}R}-c_{\tilde{R}}\\ p_{\tilde{R}}(T)=0\end{cases}\tag{5-30}$$

$$\begin{cases}\dfrac{\mathrm{d}p_D(t)}{\mathrm{d}t}=-\dfrac{\partial}{\partial D(t)}H(t,\boldsymbol{p}(t),\boldsymbol{x}^*(t),\mu^*(t),\nu^*(t))\\ \qquad =-c_D\\ p_D(T)=c_D^{\mathrm{T}}\end{cases}\tag{5-31}$$

定理 5-2 对无线传感器网络恶意程序防御微分博弈$\mathcal{G}$，无线传感器网络系统和恶意程序的最优控制分别为

$$\gamma_{IR}^{S*}(t)=\begin{cases}1, & 若\quad \eta_{IR}(t)<0\\ 0, & \quad\ 其他\end{cases}\tag{5-32}$$

$$\gamma_{I\tilde{I}}^{S*}(t)=\begin{cases}1, & 若\quad \eta_{I\tilde{I}}(t)<0\\ 0, & \quad\ 其他\end{cases}\tag{5-33}$$

$$\gamma_{SI}^{M*}(t)=\begin{cases}1, & 若 \quad \eta_{SI}(t)>0\\ 0, & 其他\end{cases} \tag{5-34}$$

$$\gamma_{ID}^{M*}(t)=\begin{cases}1, & 若 \quad \eta_{ID}(t)>0\\ 0, & 其他\end{cases} \tag{5-35}$$

证明 由式(5-24),哈密尔顿等式

$$H(t,\boldsymbol{p}(t),\boldsymbol{x}(t),\mu(t),\nu(t))=\varphi(t)+\gamma_{IR}^{S}(t)\eta_{IR}(t)+\gamma_{I\tilde{I}}^{S}(t)\eta_{I\tilde{I}}(t)$$
$$+\gamma_{SI}^{M}(t)\eta_{SI}(t)+\gamma_{ID}^{M}(t)\eta_{ID}(t) \tag{5-36}$$

是所有控制参数 $\gamma_{IR}^{S}(t)$、$\gamma_{I\tilde{I}}^{S}(t)$、$\gamma_{SI}^{M}(t)$ 和 $\gamma_{ID}^{M}(t)$ 的线性表达式。再根据式(5-22),无线传感器网络恶意程序防御微分博弈的鞍点策略($\mu^{*}(t)$,$\nu^{*}(t)$)必须满足

$$(\mu^{*}(t),\nu^{*}(t))\in\arg\min_{\mu(t)}\max_{\nu(t)}H(t,\boldsymbol{p}(t),\boldsymbol{x}(t),\mu(t),\nu(t)) \tag{5-37}$$

和

$$(\mu^{*}(t),\nu^{*}(t))\in\arg\max_{\nu(t)}\min_{\mu(t)}H(t,\boldsymbol{p}(t),\boldsymbol{x}(t),\mu(t),\nu(t)) \tag{5-38}$$

因此,定理 5-2 结论成立。证毕。

定理 5-2 说明在应用无线传感器网络恶意程序防御微分博弈时,不需要直接去计算该模型的鞍点控制策略,就可为无线传感器网络系统和恶意程序得到最优的控制。实际上鞍点控制策略的计算过程对资源有限的传感器节点而言计算太复杂而不适合传感器节点环境,而本章得到的结果实质属于 Bang-Bang 控制,这种控制模式不但能在传感器节点环境中方便地实现,而且,一旦状态函数和协状态函数已知的话,$\eta_{IR}(t)$、$\eta_{I\tilde{I}}(t)$、$\eta_{SI}(t)$和 $\eta_{ID}(t)$将会被唯一确定,从而可以方便地得到无线传感器网络系统和恶意程序的最优控制。下面给出相应的算法。

算法 5-1 计算无线传感器网络系统和恶意程序的最优控制策略。

输入:$\mathbb{G}$

输出:($\mathcal{U}^{*}$,$\mathcal{V}^{*}$)

1. 初使化所有系数。
2. 设置 $\gamma[16]$={((0,0),(0,0)),((0,0),(0,1)),…,((1,1),(1,1))}。
3. 设置 $\eta[16]$={('>0','>0','<0','<0'),('>0','>0','<0','>0'),…,('<0','<0','>0','>0')}。
4. 设置(($\gamma_{SR}^{S}(0)$,$\gamma_{I\tilde{I}}^{S}(0)$),($\gamma_{SI}^{M}(0)$,$\gamma_{ID}^{M}(0)$))=((1,1),(1,1))。
5. FOR t=1 TO T
6. 　将 $\gamma_{IR}^{S}(t)$,$\gamma_{I\tilde{I}}^{S}(t)$,$\gamma_{SI}^{M}(t)$,$\gamma_{ID}^{M}(t)$代入式(5-7)至式(5-13)和式(5-25)至式(5-31)。
7. 　采用标准的数值计算法计算式(5-7)至式(5-13)和式(5-25)至式(5-31)得到所有当前状态和协状态值。
8. 　计算($\eta_{IR}(t)$,$\eta_{I\tilde{I}}(t)$,$\eta_{SI}(t)$,$\eta_{ID}(t)$)。
9. 　FOR k=1 TO 16
10. 　　IF ($\eta_{IR}(t)$,$\eta_{I\tilde{I}}(t)$,$\eta_{SI}(t)$,$\eta_{ID}(t)$)==$\eta[k]$ THEN
11. 　　　设置($\mu^{*}(t)$,$\nu^{*}(t)$)=$\gamma[k]$。
12. 　　　Break;
13. 　　ENDIF

14.　　ENDFOR
15.　　设置$((\gamma_{SR}^{S}(t),\gamma_{I\tilde{I}}^{S}(t)),(\gamma_{SI}^{M}(t),\gamma_{ID}^{M}(t)))=\gamma[k]$。
16.　ENDFOR
17.　RETURN $(\mathcal{U}^*,\mathcal{V}^*)=\{(\mu^*(1),\nu^*(1)),(\mu^*(2),\nu^*(2)),\cdots,(\mu^*(T),\nu^*(T))\}$

5.5　实验

使用 MATLAB R2010a 实验环境，本章首先说明无线传感器网络恶意程序传播过程中各种状态的演化曲线，揭示包括由无线传感器网络系统控制的 $\gamma_{IR}^{S}(t)$ 和 $\gamma_{I\tilde{I}}^{S}(t)$ 以及由恶意程序控制的 $\gamma_{SI}^{M}(t)$ 和 $\gamma_{ID}^{M}(t)$ 等因素如何影响被感染传感器节点的数量。接下来分别说明无线传感器网络系统和恶意程序在它们交互过程中的最优控制策略。然后，将无线传感器网络系统和恶意程序都采用静态控制策略时产生的成本与它们都采用优化控制策略时产生的成本进行比较。最后给出最优控制策略对恶意程序传播影响的评价。为了完成这些实验，本章初始化必需的参数值如下：$\sigma=0.1$，$r=10\text{m}$，$N=1000$，$\gamma_{S\tilde{S}}=0.2$，$\gamma_{SI}=0.1$，$\gamma_{IR}=0.1$，$\gamma_{SD}=0.005$，$\gamma_{I\tilde{I}}=0.2$，$\gamma_{SR}=\gamma_{IR}$，$\gamma_{ID}=0.05$，$\gamma_{R\tilde{R}}=0.2$，$\gamma_{RD}=\gamma_{SD}$，$\gamma_{\tilde{S}S}=0.25$，$\gamma_{\tilde{I}I}=\gamma_{\tilde{S}S}$，$\gamma_{\tilde{R}R}=\gamma_{\tilde{S}S}$ 和 $I(0)=10$。

5.5.1　静态控制策略下各状态传感器节点数量的演化

图 5-2 和图 5-3 给出了各状态传感器节点根据时间变化的数量变化趋势。因为考虑的是静态控制策略，所以令 $\gamma_{IR}^{S}(t)$、$\gamma_{I\tilde{I}}^{S}(t)$、$\gamma_{SI}^{M}(t)$ 和 $\gamma_{ID}^{M}(t)$ 的值均为 1，这意味着无线传感器网络系统和恶意程序都尽各自最大的主观努力程度。在图 5-2 中，可以看到易感传感器节点的数量从一开始就快速下降且在时刻 30 之后下降速度变得非常缓慢。同时，在无线传感器网络系统尽最大努力安装安全补丁的前提下，康复传感器节点的数量一直在缓慢增长。由于恶意程序故意尽最大努力杀死被感染传感器节点，所以死亡传感器节点数量缓慢增长。当然，到最终将因为所有传感器节点能量耗尽而达到所有传感器节点数量的总和，此时也意味着所有除状态 D 以外其他状态的传感器节点数量都为 0。在图 5-3 中，可以看到处于休

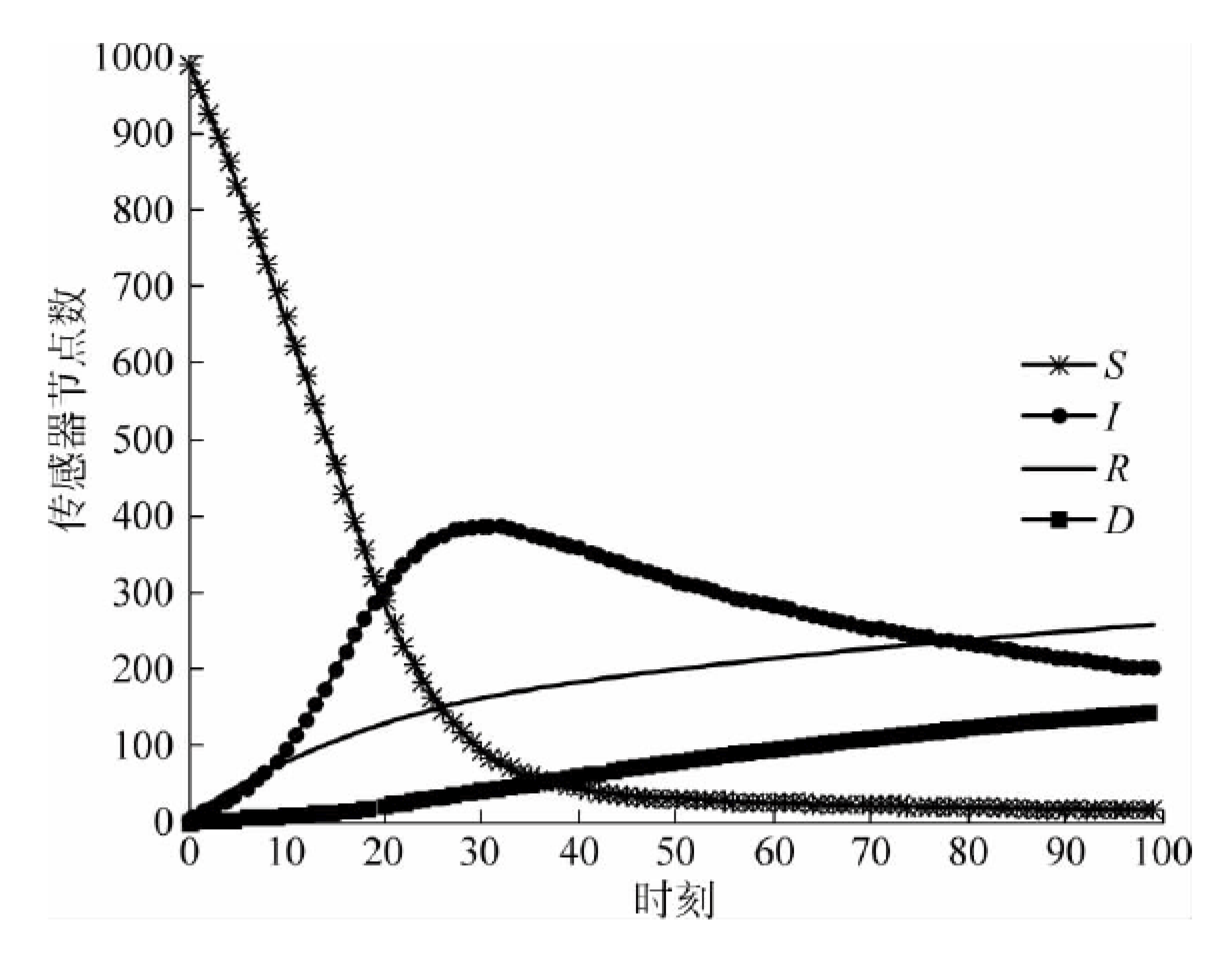


图 5-2　静态控制策略下状态 S、I、R 和 D 传感器节点数量变化趋势

眠状态的各种状态的传感器节点数量变化趋势。这些趋势有一个共同的特点，那就是所有处于休眠状态的传感器节点数量都是先增长然后再下降，然而不同状态的传感器节点数量达到最大值的时间点不同。可以预见，虽然状态$\widetilde{R}$的传感器节点数量下降缓慢，但最终所有这些处于休眠状态的传感器节点数量都将趋于0。

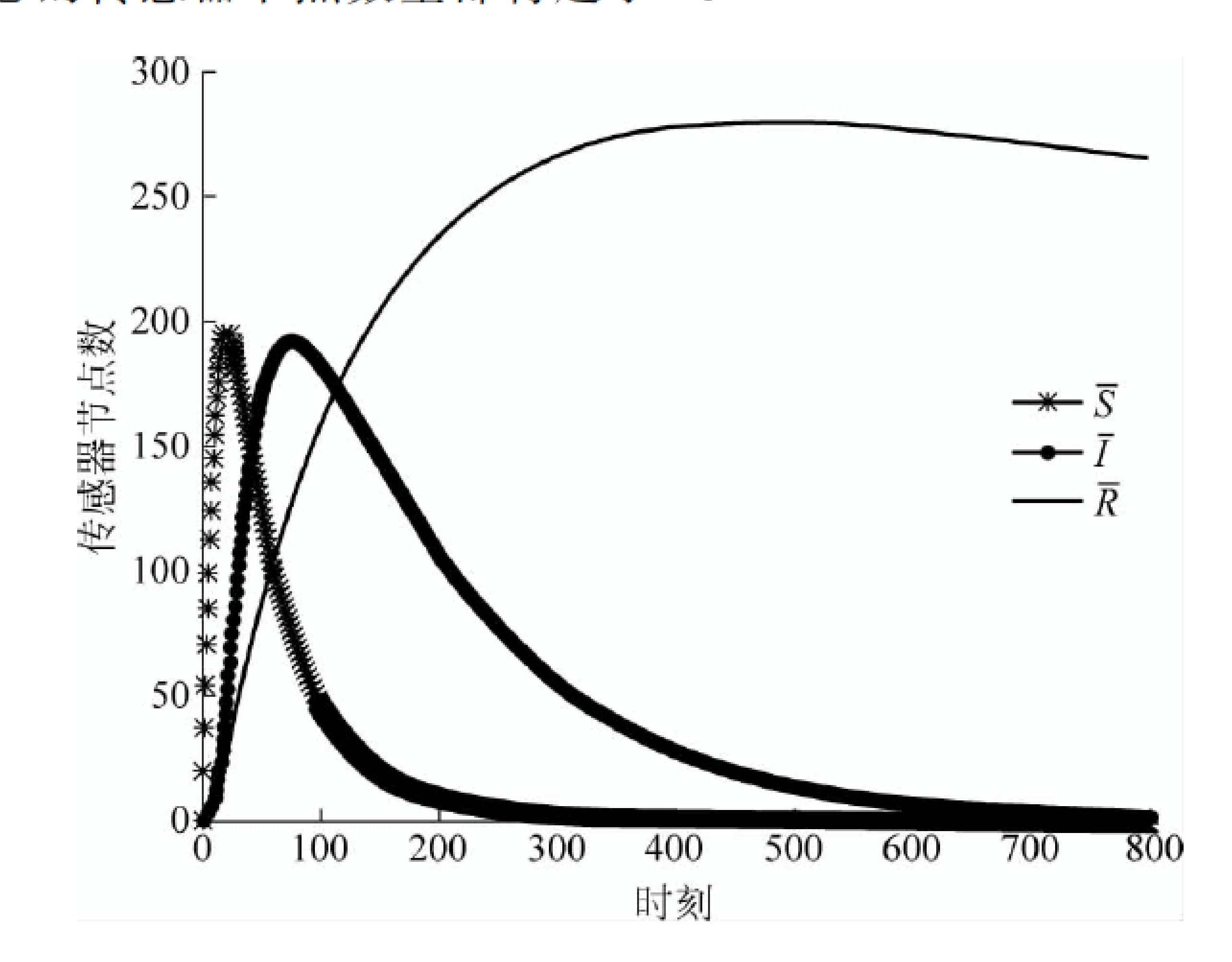


图 5-3 静态控制策略下状态$\widetilde{S}$、$\widetilde{I}$和$\widetilde{R}$传感器节点数量变化趋势

5.5.2 动态控制策略对被感染传感器节点数量的影响

首先来观察无线传感器网络系统的主观努力程度如何影响被感染传感器节点的数量。为了完成本实验，假设$\gamma_{SI}^{M}(t)$和$\gamma_{ID}^{M}(t)$是静态的且均设置为1，也就是说，恶意程序尽最大的主观努力程度，然后让$\gamma_{IR}^{S}(t)$和$\gamma_{TI}^{S}(t)$的值在区间[0,1]变化，再计算出时刻47处被感染传感器节点的数量，其变化趋势如图5-4所示。从中可以看出，随着$\gamma_{IR}^{S}(t)$的值逐渐从0变化到1，被感染传感器节点的数量缓慢下降。这个实验结果反映出无线传感器网络系统增加安装安全补丁的主观努力程度对被感染传感器节点数量的变化影响不大。显然，当$\gamma_{TI}^{S}(t)$的值从0变化到0.1时，被感染传感器节点的数量急剧下降，然而这种趋势没有持续下去，当$\gamma_{TI}^{S}(t)$值从0.1变化到1时，下降趋势变得非常缓慢。例如，当$\gamma_{IR}^{S}(t)=0$且$\gamma_{TI}^{S}(t)=0$时，$I(t)=337$，而当$\gamma_{IR}^{S}(t)=0$且$\gamma_{TI}^{S}(t)=0.1$时，$I(t)=265$，可以看到被感染传感器节点数量下降了21.36%。另外，当$\gamma_{IR}^{S}(t)=0$且$\gamma_{TI}^{S}(t)=1$时，$I(t)=248$，可见，随着$\gamma_{TI}^{S}(t)$的值从1降到0.1时，被感染传感器节点数量仅下降了6.42%。

下面来观察恶意程序的主观努力程度对传播其自身的影响，因此假设$\gamma_{IR}^{S}(t)$和$\gamma_{TI}^{S}(t)$是静态的且值均为1，这意味着无线传感器网络系统尽最大的主观努力程度，然后让$\gamma_{SI}^{M}(t)$和$\gamma_{ID}^{M}(t)$的值在区间[0, 1]变化，再计算出时刻52处被感染传感器节点的数量，其变化趋势如图5-5所示。从中可以看到随着$\gamma_{ID}^{M}(t)$从0变化到1，被感染传感器节点的数量缓慢下降。这个实验结果反映出恶意程序将状态I的传感器节点转换为状态D的主观努力程度对被感染传感器节点的数量影响有限，即使在无线传感器网络系统尽最大努力安装安全补丁时也是如此。而当$\gamma_{SI}^{M}(t)$的值从0变化到1时，被感染的传感器节点数量的增长速度很快。例如，当$\gamma_{SI}^{M}(t)=1$且$\gamma_{ID}^{M}(t)=0$时，$I(t)=381$，而当$\gamma_{SI}^{M}(t)=1$且$\gamma_{ID}^{M}(t)=1$时，$I(t)=$

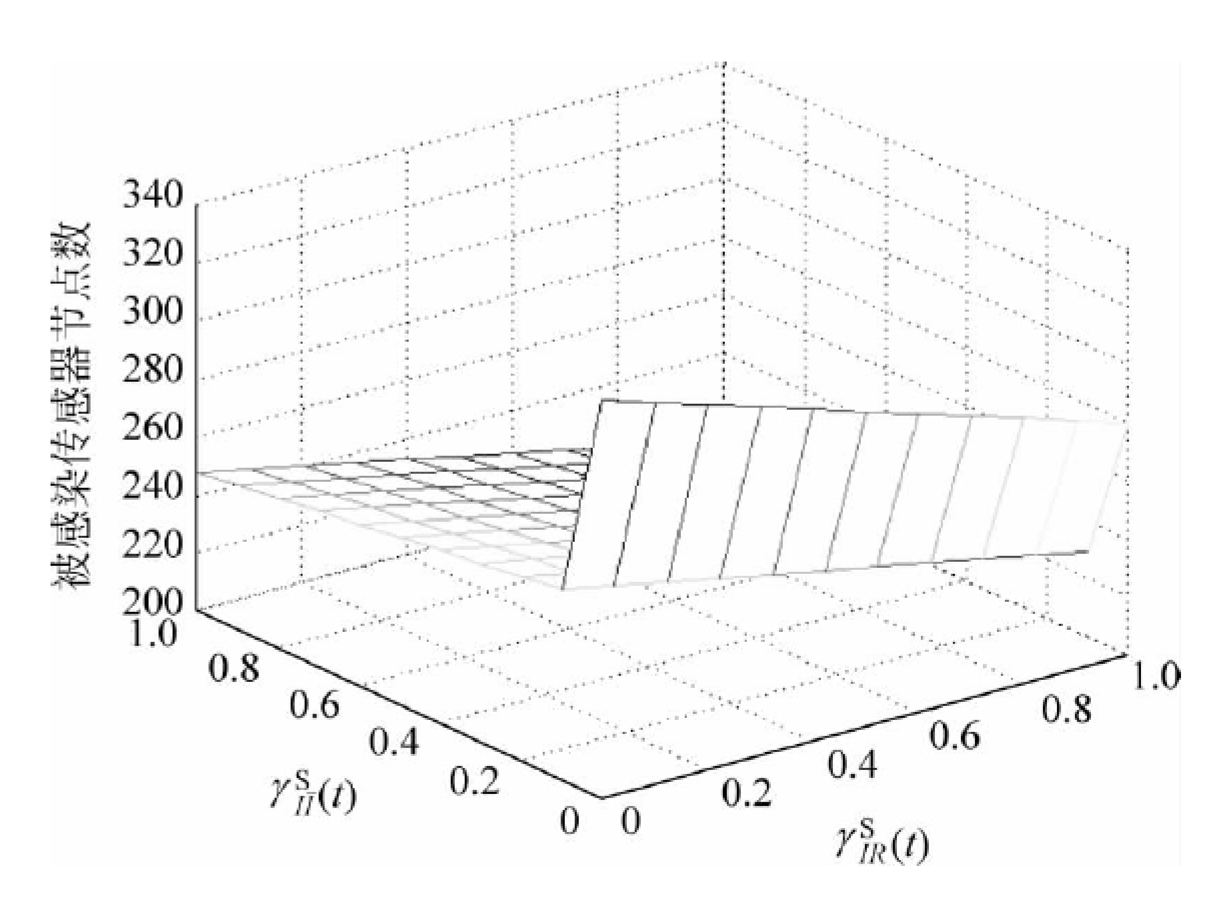


图 5-4　控制 $\gamma_{IR}^{S}(t)$ 和 $\gamma_{II}^{S}(t)$ 动态变化下的被感染传感器节点数量变化趋势

358，可见被感染传感器节点的数量仅减少 6.04%。另外，当 $\gamma_{SI}^{M}(t)=0$ 且 $\gamma_{ID}^{M}(t)=0$ 时，$I(t)=5$，这说明当 $\gamma_{SI}^{M}(t)$ 从 0 变化到 1 时，被感染的传感器节点的数量增长了 75.2 倍。

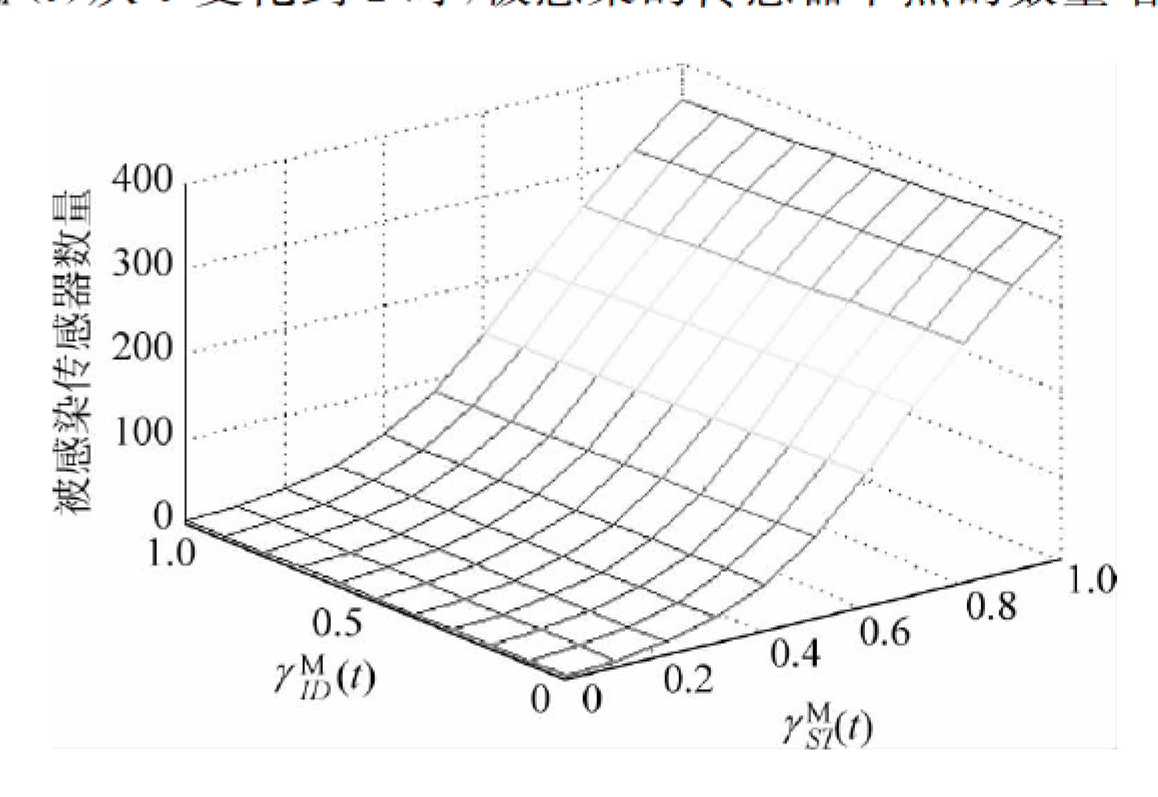


图 5-5　控制 $\gamma_{SI}^{M}(t)$ 和 $\gamma_{ID}^{M}(t)$ 动态变化下的被感染传感器节点数量变化趋势

5.5.3　无线传感器网络系统和恶意程序的最优控制策略

本实验假设整个无线传感器网络共包含 1000 个传感器节点，其中在开始时被感染的传感器节点占有很小的比例，即 $I(0)=10$。由算法 5-1 分别得到的无线传感器网络系统和恶意程序的最优控制变化如图 5-6 所示。从控制 $\gamma_{IR}^{S}(t)$ 的变化来看，无线传感器网络系统在开始时，尽最大的主观努力程度通过安装安全补丁的方式修复易感或被感染的传感器节点，但为了节省较低的带宽资源，在时刻区间(14，17)时，无线传感器网络系统停止了修复工作，而当过了时刻 17 之后，又开始了修复工作。从控制 $\gamma_{II}^{S}(t)$ 的变化来看，无线传感器网络系统在开始时也是尽最大主观努力程度去隔离被感染的传感器节点，然后在时刻 23 之后，为了保持正常的通信而停止了隔离工作，而当过了时刻 60 之后，无线传感器网络系统又开始了隔离工作。从控制 $\gamma_{SI}^{M}(t)$ 的变化来看，恶意程序从一开始不管是否会被无线传感器网络系统捕获，就尽最大努力感染那些相邻的易感传感器节点，然后在时刻 22 之后，因为被感染的传感器节点数量已达到它所期望的数值，所以停止了感染工作。从控制 $\gamma_{ID}^{M}(t)$ 的变化来看，恶意程序一开始没有努力让被感染的传感器节点失去所有功能。这里有一个有趣的

现象，那就是在时刻 24，$\gamma_{ID}^{M}(t)=1$，这意味着恶意程序应该要开始杀死被感染的传感器节点，但它毫不犹豫地停止了杀死被感染传感器节点的工作。这里的原因也许是恶意程序认为当时杀死被感染的传感器节点所获得的利益比利用被感染的传感器节点窃听私密数据所获得的利益要高，这样的行为一直持续到时刻 50，随后恶意程序开始杀死被感染的传感器节点。

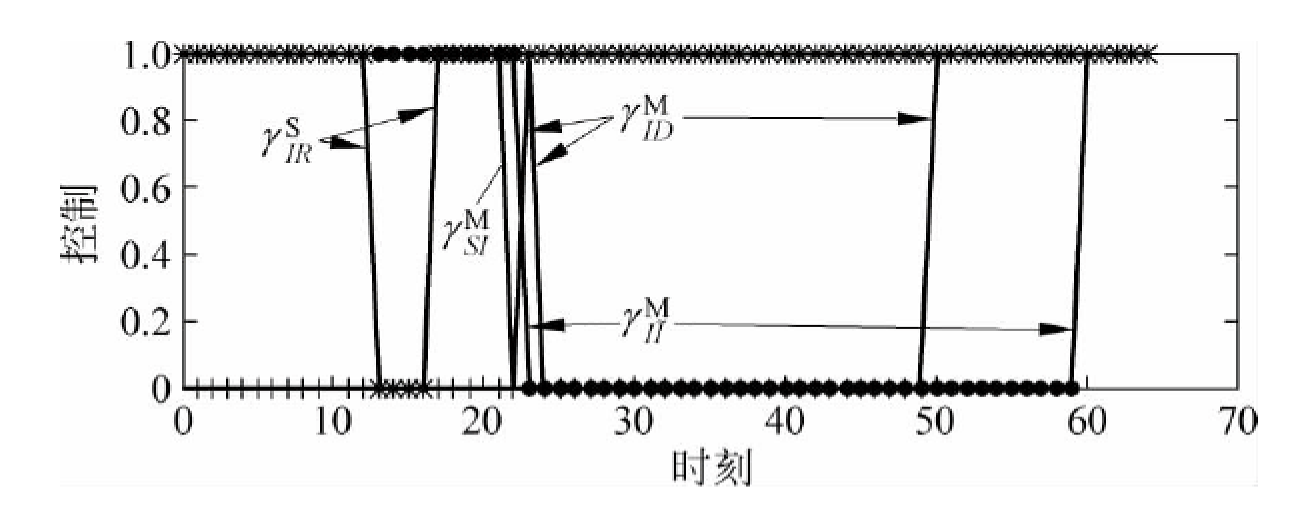


图 5-6 无线传感器网络系统和恶意程序的最优控制策略

5.5.4 静态控制策略和最优控制策略的成本比较

图 5-7 给出了式(5-15)在静态控制策略和最优控制策略下的成本比较，其中静态控制意味着无线传感器网络系统和恶意程序都尽自己最大的主观努力程度。也就是所有控制 $\gamma_{IR}^{S}(t)$、$\gamma_{II}^{S}(t)$、$\gamma_{SI}^{M}(t)$和 $\gamma_{ID}^{M}(t)$的值始终保持不变且都为 1；而最优控制策略是基于无线传感器网络恶意程序防御微分博弈，并通过算法 5-1 计算得到的无线传感器网络系统和恶意程序的鞍点策略。从图 5-7 中可以看出，最优控制策略产生的总成本小于静态控制策略产生的总成本，并且在时刻 20 后这种差异变得越来越明显，使总成本降低了 13.08%～19.94%。这些结果反映出无线传感器网络系统使用最优控制的实用性。

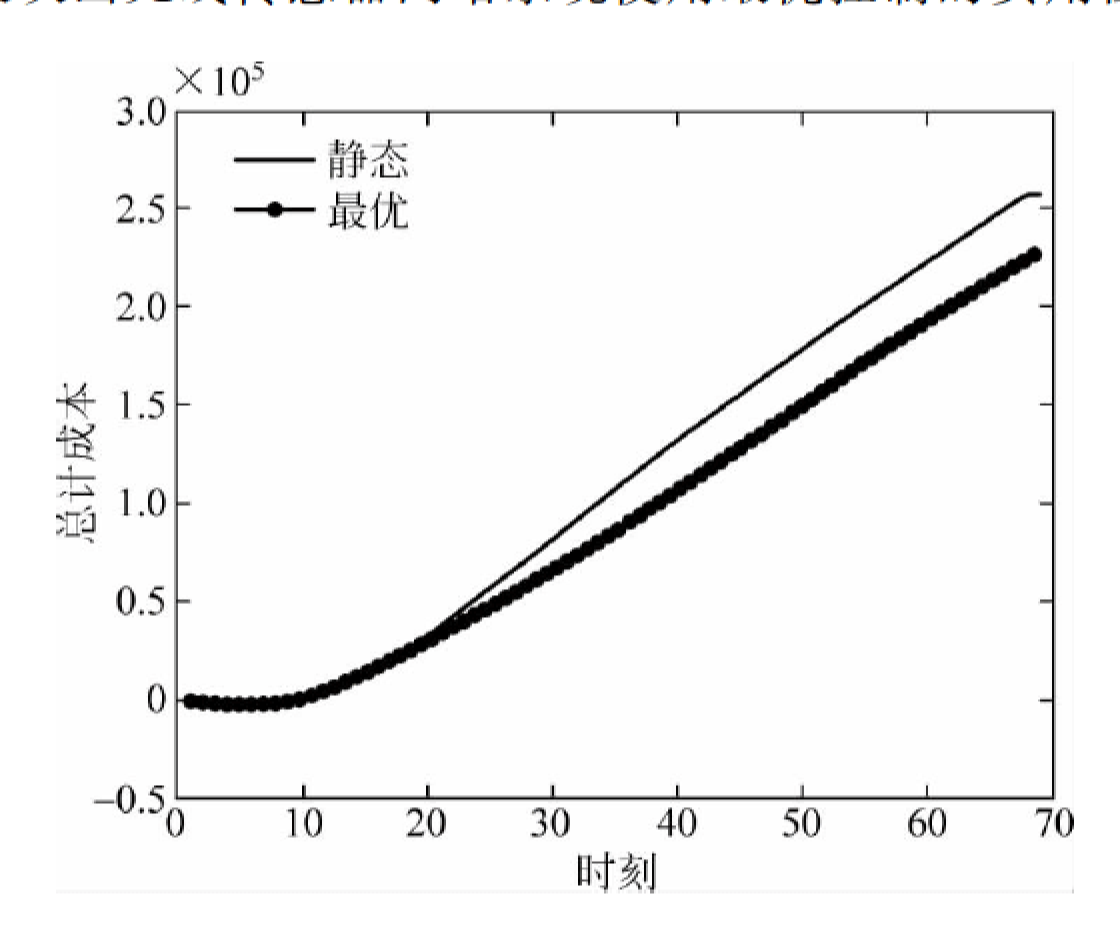


图 5-7 静态控制策略和最优控制策略下的成本比较

5.5.5 最优控制策略下的各状态传感器节点数量变化趋势

图 5-8 和图 5-9 给出了无线传感器网络系统和恶意程序在都使用最优控制下的各状态传感器节点数量的变化趋势，这些变化趋势虽与图 5-2 和图 5-3 类似，但变得更加平坦。从

中可以看出，易感传感器节点数的减少量比图 5-2 要小。因此更多的易感传感器节点为了节省能量受无线传感器网络系统调度进入休眠状态，所以状态$\widetilde{S}$中的传感器节点明显增加。而更少的易感传感器节点被恶意程序感染，因此被恶意程序故意杀死的传感器节点数明显下降。另外，除去被无线传感器网络系统隔离的易感传感器节点，进入休眠状态的易感传感器节点数也明显下降，只有康复传感器节点的数量没有明显改变。例如，在最优控制策略下，当时刻为 65 时，$S(t)=325$，$\widetilde{S}(t)=266$，$I(t)=23$，$\widetilde{I}(t)=2$，$R(t)=225$，$\widetilde{R}(t)=114$ 和 $D(t)=26$；而在静态控制下，$S(t)=23$，$\widetilde{S}(t)=94$，$I(t)=269$，$\widetilde{I}(t)=188$，$R(t)=218$，$\widetilde{R}(t)=109$ 和 $D(t)=98$。这些实验结果反映出采用最优控制能明显降低被感染和死亡传感器节点的数量，从而可以有更多的传感器节点进行正常且安全的通信，有效地延长了整个无线传感器网络的生存期。

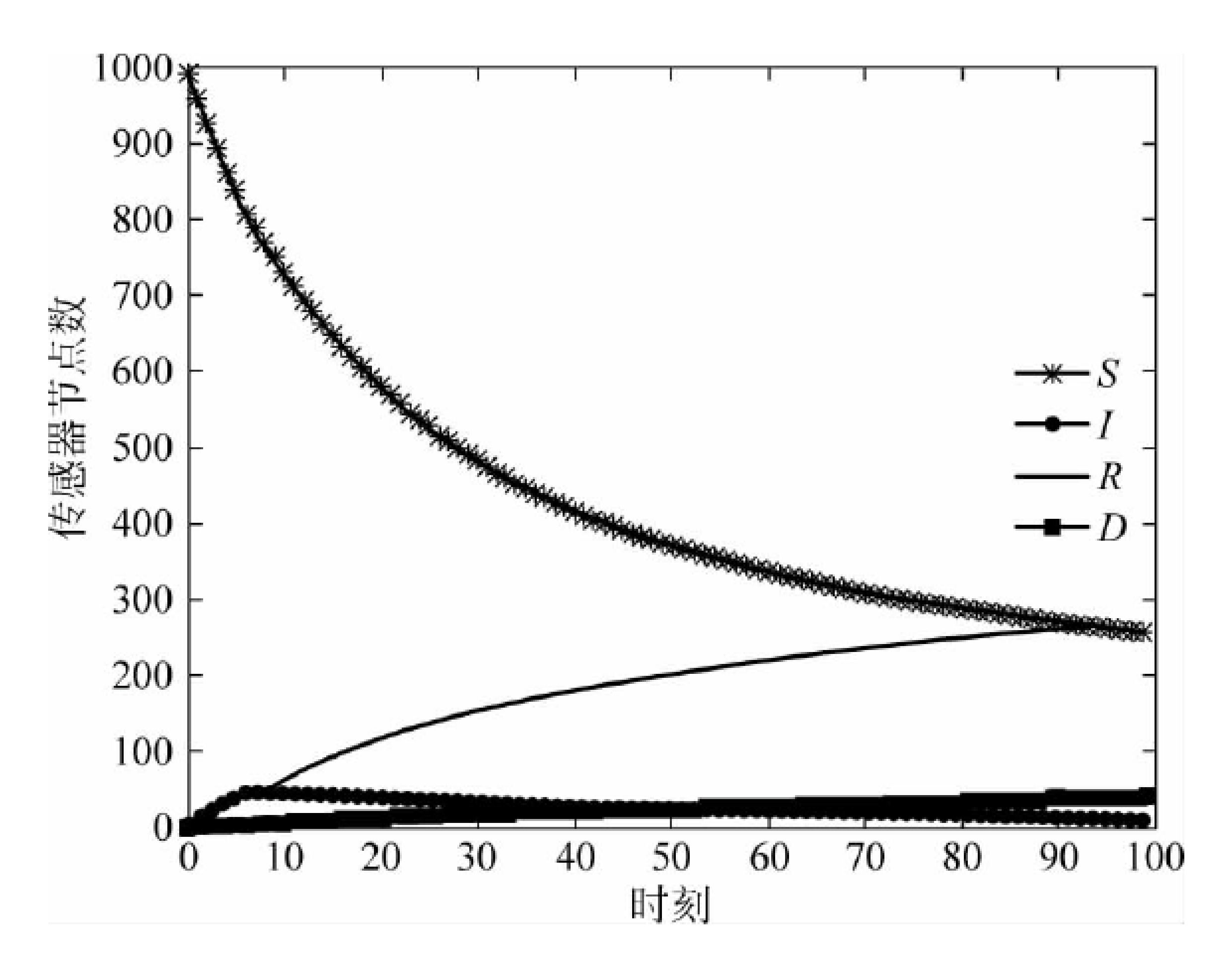


图 5-8　最优控制策略下状态 S、I、R 和 D 传感器节点数变化趋势

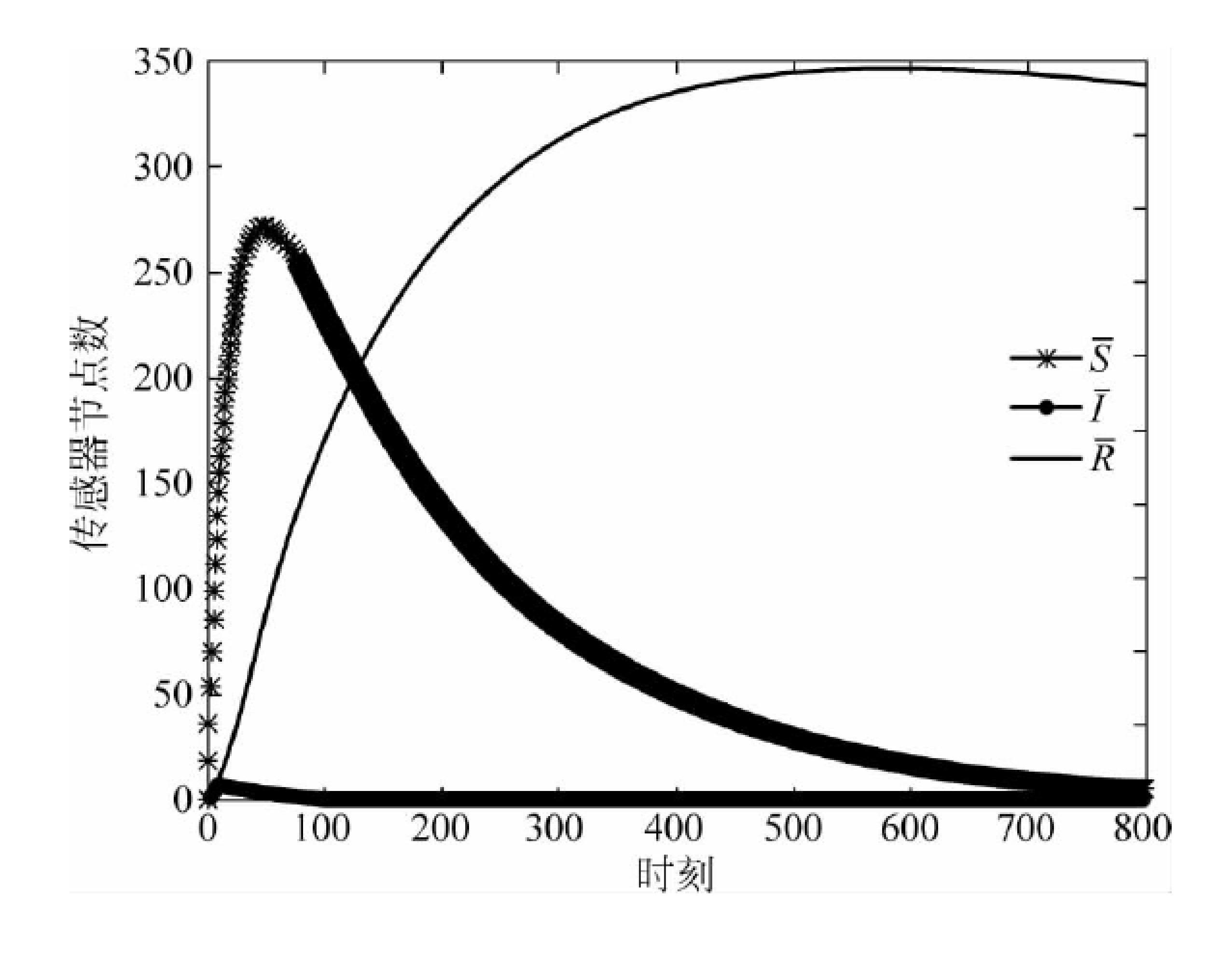


图 5-9　最优控制策略下状态$\widetilde{S}$、$\widetilde{I}$ 和$\widetilde{R}$传感器节点数变化趋势

5.6 小结

本章利用微分博弈提出了一种防御无线传感器网络恶意程序传播的方法。在考虑无线传感器网络系统和恶意程序主观努力程度的前提下，通过扩展经典的流行病理论，建立了能体现传感器节点特性的无线传感器网络恶意程序传播模型。当恶意程序动态地改变其控制策略时，利用建立的无线传感器网络恶意程序防御微分博弈能为无线传感器网络系统提供最优的控制策略，从而解决了在最大程度保证正常的网络通信与最低程度降低安装安全补丁对网络的影响之间的矛盾问题。本章得到的无线传感器网络系统最优控制策略实质上是一种 Bang-Bang 控制，因其应用方便而非常适合于传感器节点环境。实验结果验证了最优控制策略能明显地抑制无线传感器网络恶意程序的传播，体现了最优控制策略的有效性，从而有效地延长了整个无线传感器网络的生命期。

第 6 章

基于随机博弈的受攻击无线传感器网络可生存性评估研究

本章从可靠度和可用度两方面评估受攻击无线传感器网络的可生存性属性。由于恶意攻击者总是故意发动恶意攻击行为，通过随机博弈给出这些理性恶意攻击者采取恶意攻击的期望概率，将聚簇无线传感器网络看作一个串—并系统，再利用连续时间马尔可夫链对受攻击传感器节点生命期的所有状态建立模型，基于可靠性理论得到计算受攻击传感器节点平均无故障时间、可靠度、生存期和稳态可用度的计算公式，实现受攻击无线传感器网络的可生存性评估。

6.1 引言

无线传感器网络在军事、健康监测、车辆跟踪等众多领域具有广泛的用途，为了保证这些应用的顺利实现，要求无线传感器网络具备可靠和可用的能力，甚至在传感器节点被恶意攻击时仍能正常地支持具体的应用。这种能力实际上就是无线传感器网络可生存(Surviability)能力，它是无线传感器网络在安全方面的终极目标。在应用可生存性技术之前，实现可生存性评估是关键，因此，无线传感器网络可生存性评估已成为当前研究者的热点话题之一，这将为构建高可生存的无线传感器网络提供理论基础。

可生存性概念最早来源于军事通信网络，要求即使有一些通信节点在被破坏的前提下，整个军事通信网络仍能可靠地运行。它随着通信网络向现代计算网络的转变也在发生着变化，但通常认为可生存性代表了一个系统在及时完成某项任务时具有的能力，尤其在出现包括攻击和大规模自然灾害等危险的情况下[250]。从评估的观点来看，可生存性评估包括可靠度(Reliability)、可用度(Availability)和容错度(Fault-tolerance)等属性，而这些属性通常使用平均无故障时间(Mean Time To Failure)、平均修复时间(Mean Time To Repair)、平均故障间隔时间(Mean Time Between Failure)、故障率(Failure Rate)、修复率(Repair Rate)和错误覆盖率(Falut-coverage Rate)等指标体现[251]。

在可生存性评估的模型建立方面，使用连续时间马尔可夫链(Continuous-Time Markov Chain)这种建立在状态空间上的随机模型具有很大的方便性。其中的状态空间是一个由一组离散的状态组成的集合，且在任意的一个时间点上，系统精确地处于某一状态。这些状态之间的转换常使用一个状态转移矩阵进行描述，其元素代表了不同状态之间的转

移率。尤其要说明的是，连续时间马尔可夫链具备马尔可夫属性(Markov Property)，意思是转移到时刻 $t+1$ 对应状态的转移率仅依据时刻 t 的状态信息，与时刻 t 之前的状态信息无关。

然而，仅使用连续时间马尔可夫链还无法面对因恶意攻击者通过恶意攻击使整个无线传感器网络出现故障的情形，这是因为无线传感器网络中的恶意攻击者经常有意发动入侵攻击，而这种恶意攻击行为就不能使用随机过程进行描述。实际上，这个问题可考虑使用博弈论方法来解决。作为解决参与者之间决策问题的博弈论已广泛应用于网络安全领域，在各种不同的博弈模型中，随机博弈已作为一种有效的工具被用于预测攻击者恶意行为的建模[112]。通过建立随机博弈模型，可计算出一个理性的恶意攻击者采取恶意攻击的期望概率，这样，就可以把恶意攻击者成功实施攻击的概率与连续时间马尔可夫链中的状态转移率结合起来确定一个传感器节点的状态变化。另外，在恶意攻击者攻击无线传感器网络使得传感器节点状态变化的过程中，使用随机博弈不仅可以考虑恶意攻击者在正常实施攻击后得到的正收益，也可以考虑恶意攻击者被检测到后对其产生的负收益。

本章从可靠度和可用度两方面评估受攻击无线传感器网络的可生存性属性。首先将选择研究的聚簇无线传感器网络看作一个串—并系统，这样就可以应用经典可靠理论中已有的结论。因为恶意攻击者总是故意发动恶意攻击行为，通过随机博弈给出这些理性恶意攻击者采取恶意攻击的期望概率，再利用连续时间马尔可夫链对受攻击传感器节点生命期的所有状态建立模型，就可得到计算受攻击传感器节点平均无故障时间、可靠度、生存期(Survival Lifetime)和稳态可用度的计算公式，实现受攻击无线传感器网络的可生存性评估。

在扩展作者前期工作[252]的基础上，本章的工作主要包括：

(1) 在恶意攻击者和无线传感器网络系统之间建立一个零和两人攻击预测随机博弈模型，该模型能得到理性恶意攻击者在不同的传感器节点状态中的攻击概率，从而为恶意攻击者的故意攻击行为和连续时间马尔可夫链的随机性之间建立联系。

(2) 利用连续时间马尔可夫链建立受攻击传感器节点的生命期模型，该模型能描述一个传感器节点在被攻击的情况下所导致的不同状态，从而可以得到计算受攻击传感器节点的平均无故障时间。

(3) 构建受攻击无线传感器网络的可生存性评估机制，包括可靠度、生存期及稳态可用度，从而为设计高可生存的无线传感器网络奠定了理论基础。

本章其余章节安排如下：6.2 节介绍相关工作；6.3 节讨论要研究的聚簇无线传感器网络模型并把它看成是一个串—并系统，然后给出针对受攻击无线传感器网络的攻击预测随机博弈模型并说明如何预测攻击的期望概率；6.4 节从可靠度、生存期、稳态可用度 3 个方面给出受攻击无线传感器网络的可生存性评估机制；6.5 节通过实验说明攻击者采取攻击的期望概率与博弈参数之间的关系，以及这种期望概率如何影响一个受攻击传感器节点的平均无故障时间。另外，还验证了提出的可生存性评估机制的有效性；6.6 节给出本章小结。

本章涉及的符号含义如下：

Γ_k 表示整个随机博弈中的第 k 个“阶段博弈”。

z 表示整个随机博弈包含的“阶段博弈”个数。

m_k表示“阶段博弈”Γ_k 中第 1 个参与者可使用的纯策略个数。

n_k表示“阶段博弈”Γ_k 中第 2 个参与者可使用的纯策略个数。

r_{ij}^k表示当参与者 1 采取纯策略 i 且参与者 2 采取纯策略 j 时，“阶段博弈”Γ_k 的瞬时支付(Instant Payoff)。

q_{ij}^{kl}表示当参与者 1 采取纯策略 i 且参与者 2 采取纯策略 j 时，“阶段博弈”Γ_k 转换到 Γ_l 的转换概率。

μ_{ij}^k表示当参与者 1 采取纯策略 i 且参与者 2 采取纯策略 j 时，“阶段博弈”Γ_k 的累积支付(Accumulated Payoff)。

q_{ij}^{k0}表示当参与者 1 采取纯策略 i 且参与者 2 采取纯策略 j 时，“阶段博弈”Γ_k 的结束概率。

α_i^k 表示“阶段博弈”Γ_k 中参与者 1 采取纯策略 i 的概率。

β_j^k 表示“阶段博弈”Γ_k 中参与者 2 采取纯策略 j 的概率。

α^k 表示“阶段博弈”Γ_k 中参与者 1 的混合策略。

β^k 表示“阶段博弈”Γ_k 中参与者 2 的混合策略。

Δ_k 表示用于计算期望收益的矩阵博弈(Matrix Game)。

ν_k 表示初始“阶段博弈”为 Γ_k 时整个随机博弈的期望收益。

$\boldsymbol{v}$ 表示期望收益向量。

ν_{ij}^k表示当“阶段博弈”Γ_k 被其博弈值代替后的累积支付值。

$\mathbb{G}$ 表示适用于受攻击无线传感器网络的攻击预测随机博弈。

a 表示恶意攻击者采取动作 Attack，即实施攻击行为。

ϕ 表示恶意攻击者采取动作 Non-attack(即表现出正常行为)或无线传感器网络系统采取动作 Non-defend(即入侵检测机制处于关闭状态)。

d 表示无线传感器网络系统采取动作 Defend(即入侵检测机制处于开启状态)。

θ_{ij}表示恶意攻击者将阶段博弈 Γ_i 转换到 Γ_j 的恶意的努力程度。

ρ_{ij}表示无线传感器网络系统将阶段博弈 Γ_i 转换到 Γ_j 的积极的努力程度。

η 表示一个传感器节点的偶然硬件故障率(Accidental Hardware Failure Rate)。

p_{ij}表示连续时间马尔可夫链中从状态 i 转换到 j 的转移概率。

α^{k^*} 表示恶意攻击者在阶段博弈 Γ_k 的最优混合策略。

α^* 表示恶意攻击者的最优混合策略集合。

S 表示连续时间马尔可夫链中的离散状态空间。

$X_i(t)$表示一个传感器节点处于连续时间马尔可夫链中状态 i 的概率。

X_i 表示连续马尔可夫链达到稳态时状态 i 的概率。

$\boldsymbol{X}$ 表示连续马尔可夫链的稳态矩阵。

λ 表示一个传感器节点的故障率。

$R_i(t)$表示一个传感器节点的可靠度。

$R_{C_i}(t)$表示第 i 个聚簇的可靠度。

$R_{R_i}(t)$表示第 i 条路由的可靠度。

$R(t)$表示整个无线传感器网络的可靠度。

ST 表示整个无线传感器网络的生存期。

As_i 表示连续马尔可夫链达到稳态时一个传感器节点的可用度。

A_v 表示连续马尔可夫链达到稳态时整个无线传感器网络的可用度。

6.2 相关工作

由于攻击技术的不断增强，保护无线传感器网络完全不受攻击或破坏是不现实的，提高无线传感器网络的可生存性是目前解决故障和攻击问题的一种有效方法。实际上，可生存性作为无线传感器网络的核心目标，代表了无线传感器网络安全研究发展的新方向。

网络系统可生存性表示网络系统在遭受攻击和意外事故的情况下及时完成任务的能力[253]，也就是说，有任何不利条件下，网络系统可生存性反映了计算机通信系统能持续满足用户需求的能力。2000 年，Knight 和 Sullivan[254]给出了网络系统可生存的一种四元组表示方法。杨超和马建峰[255]较早提出了规范化的网络系统的可生存性定义，并给出网络系统可生存性的形式化描述及其实现模型。Habib 等人[256]综述了光通信网络中可生存性研究的现状，并对现有技术作了合适的分类。Albano 等人[257]综述了车联 Ad Hoc 网络中容错性、可恢复性、可生存性等研究的现状。

可生存技术是网络系统在入侵和故障已发生的情况下，仍能使网络具有完成关键任务的能力[250]。这些技术体现在无线传感器网络中，主要是冗余节点部署[258-262]、多路径路由[263-266]、多重覆盖[267-271]、操作系统容错[272]、网络编码[273-275]、入侵容忍的安全架构[276-279]等。张万松和王立松[272]从外部攻击和内部错误两方面考虑改进无线传感器网络节点操作系统 TinyOS 的生存性，设计了一种入侵检测与恢复机制，并将 TinyOS 的调度机制改进为支持容错的实时调度策略。

要对网络的可生存性属性进行量化评估，首先需要建立形式化的数学模型。目前，网络可生存性评估的建模技术主要基于系统结构[280]、状态和服务组件[281, 282]、脆弱模型[283]、概率模型[284, 285]、改进的逼近理想解排序法[286]、二项式模型[287]、模糊综合评价[288-290]、攻击树[291]、传染病模型[292, 293]、层次化评估[294, 295]等。相比之下，当无线传感器网络节点受到入侵时，基于状态的建模技术更适于描述节点的变化情况。而能对网络系统状态进行有效描述的随机模型有多种，如马尔可夫链[296-302]、马尔可夫报酬过程[303]、随机 Petri 网[304-308]、广义随机 Petri 网[309]、着色 Petri 网[310]、随机博弈网(Stochastic Game Net)[311]等，其中马尔可夫链是一种经典方法。Buzacott 很早将连续时间马尔可夫链用于寻找可修复系统的故障时间，在其经典论文[312]中利用连续时间马尔可夫链描述了可修复系统的各状态变化，并建立了平均无故障时间的计算公式。Sallhammar 等人[313]融合随机博弈和连续时间马尔可夫链用于系统的可依赖度(Dependability)评估，其中随机博弈被用于计算攻击者采取攻击的概率，从而将攻击者的攻击行为与系统之间的状态转变建立了联系，并采用与 Buzacott 相同的方法[312]将连续时间马尔可夫链用于计算平均无故障时间。Ghazisaidi 等人[284]利用概率方法评估无源光纤网络(Passive Optical Networks)的可生存性。Zhao 等人[286]利用改进的逼近理想解排序法(TOPSIS)和灰关联分析法(Grey Relation Analysis)评估网络系统的可生存性，其中逼近理想解排序法被用于指示矩阵的规范化，灰关联分析法被用于计算每个关键服务的关联度并据此得到最优的从属度(Dependency Degree)，从而实现整个网络的可

生存性评估。沈建春等人[288]采用 Delphi 方法对影响网络系统生存性的各种因素进行分析，确立了信息网络系统可生存性评价指标体系，提出了一种基于模糊数学方法的评估模型以适应信息网络系统可生存性评估的复杂性和不确定性。Zhao 和 Yu[291]利用攻击树构建入侵环境，通过计算入侵的风险度确定关键服务的可生存性。熊琦等人[296]利用随机博弈建立了入侵者和入侵容忍系统之间的随机博弈模型，提出了面向可生存性研究的容侵系统状态转换模型，使用连续马尔可夫链对容侵系统的可生存性进行了量化分析和评估。谢波等人[297]给出了满足车辆自组织网络(VANET)特点和实际应用的可生存性定义，提出了基于马尔可夫链的平均可生存性量化模型。Jindal 等人[298]利用马尔可夫链评估蜂窝网的可生存性。Wang 和 Yu[299]结合使用马尔可夫链和排队论评估 Ad Hoc 网络的可生存性。刘密霞等人[307]利用 Petri 网对分布式网络系统进行形式化描述与建模，构建了分布式网络系统的攻击失效模型，并用模糊推理方法描述分布式网络系统在攻击发生时状态的变化，提出了分布式网络系统可生存性的评价参数。刘梅霞和古天龙[309]利用马尔可夫链分析 Ad Hoc 网络的可生存性，提出了可以表示 Ad Hoc 网络中任意两个节点之间的动态数据传输关系和受故障影响情况下的广义随机 Petri 网模型。通过计算任意两个节点连通的概率，从节点传输范围、节点平均邻居数目和故障频率等方面分析 Ad Hoc 网络的可生存性。Xing 和 Wang[314]为评估 Ad Hoc 网络的可生存性，利用半马尔可夫过程，建立了描述恶意节点状态的模型。Peng 等人[315]利用连续时间马尔可夫链建立了描述各种故障状态的模型。根据文献[254]中可生存性定义，Chen 等人[316]将故障导致的过度包丢失(Excess Packet Loss due to Failure)作为 Ad Hoc 网络可生存性指标，并将文献[254]中描述系统可生存性的有限状态自动机转换为连续时间马尔可夫链从而实现过度包丢失的计算。Peng 等人[315]在考虑节点软硬件错误和连接状态基础上研究大规模移动 Ad Hoc 网络中的可生存性评估问题，采用与 Chen 等人[316]相同的方法将有限状态自动机转换为连续时间马尔可夫链，并在传统可靠性理论基础上从稳态可用度、连通度、故障节点平均数、平均生命期等方面给出了段段路由(Segment-by-segment Routing)、多路径段段路由(Multipath-based Segment-by-segment Routing)、段段多路径路由(Segment-by-segment-based Multipath Routing)的可生存性评估。Sedaghatbaf 和 Abdollahi Azgomi[317]在扩展随机活动网络(Stochastic Activity Networks)基础上提出了一种网络攻击建模方法，实现了对网络保密性、完整性、可用性等的分析，并从安全失效平均时间(Mean Time to Security Failure)和攻击成功概率(Attack Success Probability)两方面量化评估网络的安全性。

然而，当前对网络系统可生存性评估的指标并不统一。根据 Al-Kuwaiti 等人[251]的观点，可生存性包含可靠性、可用性、容错性、安全性等属性。相比较而言，较多的文献关注无线传感器网络的可靠度评估，这些方法主要有有序二叉判定图[318, 319]、增强有序二叉判定图[320]、连续 PH 分布(Continuous Phase type distributions)和 Kronecker 代数[321]、排队论模型[322]。肖坤等人[323]采用韧性度刻画网络的脆弱性，提出了一种基于韧性度的 Ad Hoc 网络可生存性度量方法。肖志力等人[324]结合联合分析法和层次分析法，提出一种综合评价方法，用于评估网络信息系统的可生存性。魏昭等人[325]提出了一种用于评判多种移动 Ad Hoc 网络可生存性模型的建模及其仿真验证方法。Kim 等人[326]研究多通道 Ad Hoc 军事网络的可生存性评估方法，给出了多通道环境下连通度的定义，通过计算节点间单跳链接的数量评估网络的可生存性。Wang 等人[327]提出了一种统一的网络可生存性评估框架，

这种框架具有可扩展和用户自定义的特点，给出了具体的测试过程。Wang 和 Yu[328]将灰关联分析法用于评估 Ad Hoc 网络的可生存性。

与 Al-Kuwaiti 等人[251]观点不同的是，Ming 等人[329]认为网络系统可生存性评估的指标应包括可用性(Availability)、可控制性(Controllability)、鲁棒性(Robustness)、适应性(Adaptability)4 个方面。Sterbenz 等人[330]则将可恢复性(Resilience)、可生存性、崩溃容忍性(Disruption Tolerance)评估网络的指标。Lin 等人[331]采用平均非连通度作为评估网络被恶意攻击后的可生存性指标。吴庆涛等人[332]认为应把数据机密度、数据完整度、服务可用度和系统自律度作为自律入侵容忍系统的可生存性评估指标。其他的指标还有 K-连通度[333]、攻击环境下的数据分发率(Delivery Rate under Attack)[334]等。Rak[335]针对无线 Mesh 网络容易产生区域失效(Region Failure)的问题，提出了区域失效可生存性函数、p 比例区域可生存性函数(P-fractile Region Survivability Function)、失效后总发送数据流期望比率(Expected Percentage of Total Flow Delivered after a Failure)等评估指标。

作为博弈论中的一种博弈类型，随机博弈已被广泛应用于与状态转移相关环境中的参与者决策问题。Lye 和 Wing[336]利用随机博弈分析传统计算机网络中恶意攻击者和网络管理者之间的交互，使用非线性规划方法计算得到纳什均衡，从而为管理者增强网络安全提供了最优响应策略。Chen 等人[337]考虑受攻击网络中的态势感知问题，通过数据融合策略实现威胁的检测和预防，其中各种威胁的检测在第二层数据融合时由智能代理实现，而预测在第三层数据融合中由一个分布式的随机博弈模型实现。Liu 等人[338]提出一种内部随机博弈(Insider Stochastic Game)解决内部威胁问题，通过预测内部攻击者的恶意行为，得到最优的防御策略。Nguyen 等人[339]在攻击者和防御者之间建立了基于随机博弈的安全博弈模型，得到的纳什均衡被用于帮助人们理解攻击者的行为，从而为入侵检测系统提供了如何防御的指导。为了给无线网络中各种用户提供一个相互成功竞争可用频谱资源的机会，Fu 和 Schaar[340]利用随机博弈分析了给定频谱干扰环境下用户之间的交互。Niyato 等人[341]基于通道保留共享方法(Channel Reservation Sharing Method)提供了一种无线通道访问模式，其中将联盟博弈用于共享保留的通道以最小化无线通道访问的成本，而当移动用户的 QoS 需求面临冲突时，随机博弈被用于协调各用户对同一个无线通道的访问。为了解决认知无线电网络中的拥塞攻击问题，Wang 等人[342]提出一种能适应环境动态变化和攻击者策略变化的反拥塞随机博弈(Anti-jamming Stochastic Game)，实现拥塞攻击的防御。

无线传感器网络可生存性分析和评估是网络可生存性理论的主要研究内容[343]。然而，当前关注无线传感器网络可生存性评估的文献还不多。Di Pietro 和 Verde[292]提出应用传染病模型研究无照料的无线传感器网络的数据可生存性。Parvin 等人[300]利用马尔可夫链建立了描述无线传感器网络节点在受攻击时各种状态变化的模型，并以 DoS 攻击为例，确定各状态进行转换的阈值。他们[344]还使用软件再生(Software Rejuvenation)技术增强无线传感器网络的可用性和可生存性，利用马尔可夫链建立了反映节点变化的状态集。Xiao 等人[320]在考虑普通传感器节点错误基础上利用增强二叉决策图算法(Enhanced Ordered Binary Decision Diagran Algorithm)评估无线传感器网络的可靠度。Korkmaz 和 Sarac[345]为了给可靠数据传输协议提供可能的设计选项。通过量化单跳无线连接的可靠度实现了整个数据转发路径的可靠度评估。何明等人[346]通过确定是否满足 K-覆盖和 K-连通来评估无线传感器网络的可靠度。在容错度评估方面，王良民等人[347]给出了拓扑容错

度和容侵度作为拓扑对节点失败容忍能力高低的评估标准。在安全度评估方面，詹永照等人[348]利用 Monte Carlo 方法评估无线传感器网络的路由安全。在可生存性评估方面，Masoum 等人[349]在考虑网络错误和这些错误对无线传感器网络影响的基础上提出无线传感器网络可生存性评估机制，这些机制包括无线传感器网络达到稳态时的网络连通度(Connectivity)和覆盖度(Coverage)。王海涛等人[350]给出了应急通信中无线传感器网络的可生存性评价指标集，建立了基于网络分析法(ANP)的评价模型框架，然后，通过该方法确定了相应指标的权重并构建了无线传感器网络生存性指标体系。朱世才等人[351]提出了一种基于半马尔可夫过程(SMP)的分簇无线传感器网络可生存性评估模型，该模型在考虑应急通信中簇头生存状态的基础上建立了基于 SMP 的簇头生存状态转移图，再结合网络生存性需求计算无线传感器网络的生存性效用函数，并定量分析了多种评价指标对网络可生存能力的影响。另外，Ma 和 Krings[352, 353]提出应用动态混合故障(Dynamic Hybrid Fault)模型和演化博弈研究无线传感器网络可生存性的思想，利用演化博弈描述节点间的相互行为，认为节点的支付就是它的可靠度，当达到演化稳定策略(Evolutionary Stable Strategy)时就能保证网络的可生存性，从而实现无线传感器网络的可生存性评估。Petridou 等人[354]从失效频率、数据丢包率、数据延迟率、数据泄露率等评估无线传感器网络的可生存性。

与上述相关工作不同的是，本章着重关注聚簇无线传感器网络的可生存性评估。本章定义的能预测理性恶意攻击者采取何种动作概率的零和双人攻击预测随机博弈类似于 Sallhammar 等人[313]的博弈模型，且本章根据 Sallhammar 等人[313]的观点建立了恶意攻击者累积的导致传感器节点故障的努力程度和连续时间马尔可夫链中状态转移率的关系。然而，本章与 Sallhammar 等人[313]明显不同的是，本章定义的攻击预测随机博弈很好地满足了聚簇无线传感器网络的特性。更进一步，本章为了给构建高可生存的无线传感器网络设计提供理论基础，因此建立了受攻击无线传感器网络的可生存性评估机制。而 Sallhammar 等人[313]主要关注基于随机模型技术的安全度和可依赖度评估。本章采用传统的连续时间马尔可夫链反映一个受攻击传感器节点的状态变化过程，给出的状态足以适合传感器节点在受到恶意攻击时的实际状况。本章虽然采用与 Buzacott[312]相同的方法计算一个受攻击传感器节点的平均无故障时间，但本章进一步得到了受攻击传感器节点的故障率和可靠度，从而可以推导出整个无线传感器网络的可靠度、生存期、稳态可用度等公式，这些公式构成了部署在受攻击环境下无线传感器网络可生存性评估的整个机制。另外，本章通过实验说明了恶意攻击者的期望动机、受攻击传感器节点的平均无故障时间以及整个无线传感器网络的可靠度、生存期和稳态可用度。

6.3 基于随机博弈的恶意传感器节点期望动机预测

6.3.1 网络模型

根据无线传感器网络结构的不同组织形式，相应的路由协议可以分成两大类：平面网络结构路由和层次网络结构路由。在平面网络结构路由中，每个传感器节点在与其他传感器节点通信时担任相同的角色，以泛洪的方式寻找一条能到达汇聚节点的路由。这种方法对规模较小的网络而言工作效率非常高，然而由于其在路由发现时需要发送大量泛洪信息

而不适合密度大的网络。另外，在层次网络结构路由中，不同的传感器节点具有不同的角色并以聚簇的形式被组织起来，这种结构已被公认能有效延长无线传感器网络的生命期。

正因为无线传感器网络聚簇结构具有的优点，本章选择该结构研究其可生存性评估方法。如图 6-1 所示，在这种聚簇无线传感器网络中，自然地可将每个簇对应一个并行系统，再将簇的集合看作是一个串行系统，所以，从一个源传感器节点到基站的路由就可看作是一个串—并系统。当然，路由数不止一条，因此，整个无线传感器网络就可以看作是一个复杂的并行系统，数据可以通过不同的并行路由进行传输。实际上，聚簇无线传感器网络与经典可靠性理论中的串—并系统本身就有很多相似之处。无线传感器网络中的每个传感器节点所处的地理环境大致相同，它们相互独立。也就是说，一个传感器节点出现故障不会导致其他传感器节点产生故障。另外，传感器节点经常会被冗余部署从而增强整个无线传感器网络工作的可靠性，这种方法在典型的串—并系统中也是被广泛应用，从而增强其可靠性。在图 6-1 中，只要每个簇的簇头因恶意攻击出现故障，无线传感器网络就会立刻从候选簇头中选举出新的簇头。因此，只要还有一个候选簇头能正常工作，那么该候选簇头所在的簇即能正常工作，这实际上正是一个并行系统所具有的特性。而当一个成员传感器节点感知到相应数据并通过其他相邻的簇头传输到基站时，可以看到在整个数据传输过程中，只有所有路由经过的簇头正常工作，数据才能被正常地传输到基站，这实际上正是一个串行系统所具有的特征。因此，一条路由中所有簇头的集合就可以看成由独立组件构成的一个串行系统。

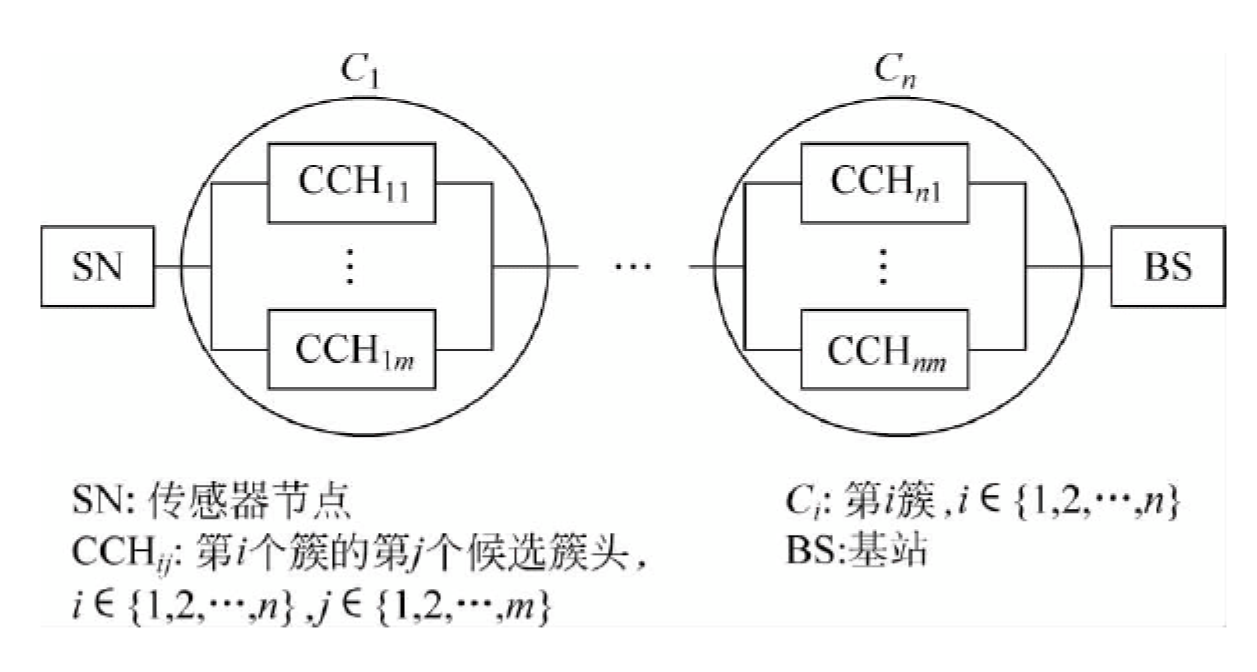


图 6-1 基于串—并行系统的聚簇无线传感器网络结构

6.3.2 无线传感器网络攻击预测随机博弈模型

定义 6-1 面向聚簇无线传感器网络的双人零和攻击预测随机博弈是一个五元组$\mathbb{G}=(N,\Gamma,A,Q,U)$，其中：

- $N=\{$恶意攻击者，无线传感器网络系统$\}$表示参与者集合。
- $\Gamma=\{\Gamma_V,\Gamma_W,\Gamma_C\}$表示"阶段博弈"集合，其中 Γ_V、Γ_W、Γ_C 分别表示受攻击传感器节点的 3 个状态：脆弱(Vulnerable)、衰弱(Weak)、妥协(Compromised)。
- $A=A_1\times A_2$ 表示恶意攻击者可采取的策略集合与无线传感器网络系统可采取策略集合的笛卡儿积，其中 $A_1=\{\text{attack, non-attack}\}$是恶意攻击者可采取的策略集合，$A_2=\{\text{defend, non-defend}\}$是无线传感器网络系统可采取的策略集合。
- $Q:\Gamma\times A\times\Gamma\mapsto[0,1]$是由各阶段博弈转移矩阵组成的集合。
- $U:\Gamma\times A_1\times A_2\mapsto\mathbb{R}$ 是由各阶段博弈的支付矩阵组成的集合。

在定义 6-1 中，考虑整个随机博弈的参与者包括恶意攻击者和无线传感器网络系统。虽然，在实际的无线传感器网络中有各种各样的攻击者，但它们的目的都是为了破坏传感器节点。因此，使用参与者恶意攻击者能代表那些行为相似的所有恶意攻击者。参与者无线传感器网络系统实际上是无线传感器网络的入侵检测机制。由于定义攻击预测随机博弈的目的主要是预测恶意攻击者的期望动机，因此在接下来的讨论中主要从参与者恶意攻击者方面说明博弈的过程。

一个传感器节点的生命期包含有限的一些状态，这些状态可以使用连续时间马尔可夫链进行描述，相应的各状态之间的转换关系如图 6-2 所示。需要注意的是图 6-2 中的 p_{ij} 表示连续时间马尔可夫链中的状态 i 转换到 j 的概率，而不是随机博弈中阶段博弈转换的概率。

在图 6-2 中，虽然一个传感器节点整个生命期的状态包括：健康(Healthy，H)、脆弱(Vulnerable，V)、衰弱(Weak，W)、妥协(Compromised，C)、失败(Failed，F)，但恶意攻击者关心的是除状态 H 和 F 外的其他状态，因此，阶段博弈集合应该是 $\Gamma=\{\Gamma_V,\Gamma_W,\Gamma_C\}$ 且整个攻击预测随机博弈从阶段博弈 Γ_V 开始。但从传感器节点而言，任何一个传感器节点在开始时处于状态 H，当无线传感器网络入侵检测机制不能成功地检测到恶意攻击者的行为且恶意攻击者通过探测已发现传感器节点存在漏洞时，传感器节点的状态即从 H 转换到 V。恶意攻击者接下来可能会利用这些漏洞并发动攻击使传感器节点状态转换为 W。为了得到更多的利益，恶意攻击者可能会持续攻击并突破传感器节点的安全防线直至将传感器节点状态转换为 C。这种被转换为状态 C 的传感器节点可能已被破坏，也可能变成恶意传感器节点，从而干扰整个无线传感器网络的通信。另外，任意状态的一个传感器节点都有可能因为偶然的软、硬件故障导致状态转换到 F。

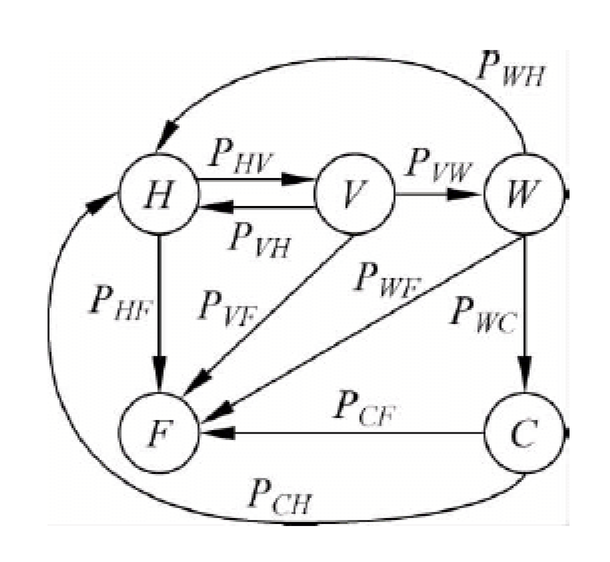


图 6-2　基于连续时间马尔可夫链的状态转换

为简化起见，恶意攻击者在所有的状态中包含两种动作，即 attack(a)和 non-attack(ϕ)，其中选择动作 a 表示实施攻击行为，选择动作 ϕ 表示伪装(即不实施任何攻击)行为。因此，恶意攻击者对应的混合策略可表示为

$$\alpha^k=(\alpha_a^k,\alpha_\phi^k) \tag{6-1}$$

其实质代表了状态 k 恶意攻击者在行动空间 A_1 上动作的概率分布，当然，满足条件

$$\alpha_a^k+\alpha_\phi^k=1 \tag{6-2}$$

实际上，α_a^k 代表了恶意攻击者对无线传感器网络的敌对程度，其值越大，在状态 k 采取攻击的概率越大，从而导致传感器节点的故障率越高。针对恶意攻击者的动作行为，无线传感器网络系统可供选择的动作行为包括 defend(d)和 non-defend(ϕ)。其中选择动作 d 表示采取防御行为，选择动作 ϕ 表示未采取任何防御行为。相应地，在状态 k 无线传感器网络系统的混合策略为

$$\beta^k=(\beta_d^k,\beta_\phi^k) \tag{6-3}$$

且满足条件

$$\beta_d^k+\beta_\phi^k=1 \tag{6-4}$$

为了计算从阶段博弈 Γ_k 转换到 Γ_l 中的转换概率 q_{ij}^{kl}，需要对恶意攻击者选择的动作和

对应的无线传感器网络系统选择的动作有一个假设前提。当整个攻击预测随机博弈中的阶段博弈从 Γ_V 转换到 Γ_W，或从 Γ_W 转换到 Γ_C 时，意味着恶意攻击者的攻击行为未被检测到且被成功实施，因此，可得到恶意攻击者和无线传感器网络系统选择的动作对为(a,ϕ)。同时，这些阶段博弈的转换还跟恶意攻击者累积的导致传感器节点故障的努力程度有关。设 θ_{ij} 为恶意攻击者将阶段博弈 Γ_i 转换到 Γ_j 累积的导致传感器节点故障的努力程度，ρ_{ij} 为无线传感器网络系统将其他的状态转换为 H 的主观努力程度，η 为一个传感器节点的偶然硬件故障率，对所有的阶段博弈 Γ_i、Γ_j，结合在阶段博弈 Γ_k 时恶意攻击者选择动作 a 的概率 α_a^k，就可以定义 θ_{ij} 和 p_{ij} 之间的关系。例如，对阶段博弈 Γ_V，可以得到

$$p_{VW} = \alpha_a^V \theta_{VW} \tag{6-5}$$

$$p_{VH} = \rho_{VH} \tag{6-6}$$

$$p_{VF} = \eta \tag{6-7}$$

因此，阶段博弈 Γ_V 转换到 Γ_W 的转换概率为

$$q_{ij}^{VW} = \begin{cases} \alpha_a^V \theta_{VW} / (\alpha_a^V \theta_{VW} + \rho_{VH} + \eta), & 若\ i = a\ 且\ j = \phi \\ 0, & 其他 \end{cases} \tag{6-8}$$

类似地，阶段博弈 Γ_W 转换到 Γ_C 的转换概率为

$$q_{ij}^{WC} = \begin{cases} \alpha_a^W \theta_{WC} / (\alpha_a^W \theta_{WC} + \rho_{WH} + \eta), & 若\ i = a\ 和\ j = \phi \\ 0, & 其他 \end{cases} \tag{6-9}$$

最后来分析攻击预测随机博弈中的支付矩阵集合。对每个阶段博弈 Γ_k，恶意攻击者与无线传感器网络系统交互时能得到一个瞬时支付 r_{ij}^k，其值若为负数则表示恶意攻击者采取动作所产生的收益小于支出的成本。由于恶意攻击者可选择的动作包括 a 和 ϕ，无线传感器网络系统可选择的动作包括 d 和 ϕ，因此总共有 4 种“动作对”，即产生 4 个瞬时支付。对“动作对”(a, ϕ)而言，它能使阶段博弈 Γ_V 转换到 Γ_W 或 Γ_W 转换到 Γ_C，这样的结果是恶意攻击者最希望看到的。因为这种结果将给恶意攻击者带来正收益；对动作对(a, d)而言，这是恶意攻击者最不希望碰到的。因为这意味着无线传感器网络系统将积极防御恶意攻击者采取的恶意攻击，从而使恶意攻击者遭受损失；其他的两个动作对(ϕ,ϕ)和(ϕ, d)对恶意攻击者而言不会产生收益或支出成本，但它们会被用于计算恶意攻击者在是否选择动作 a 或 ϕ 时的概率。由式(2-16)，就可得到阶段博弈 Γ_V、Γ_W、Γ_C 的累积支付分别为

$$\mu_{ij}^V = \begin{cases} r_{ij}^V + q_{ij}^{VW} \Gamma_W, & 若\ i = a\ 且\ j = \phi \\ r_{ij}^V, & 其他 \end{cases} \tag{6-10}$$

$$\mu_{ij}^W = \begin{cases} r_{ij}^W + q_{ij}^{WC} \Gamma_C, & 若\ i = a\ 且\ j = \phi \\ r_{ij}^W, & 其他 \end{cases} \tag{6-11}$$

$$\mu_{ij}^C = r_{ij}^C \tag{6-12}$$

6.3.3 基于攻击预测随机博弈的攻击预测算法

预测恶意攻击者的攻击行为实质是计算聚簇无线传感器网络攻击预测随机博弈中恶意攻击者的最优策略，要注意的是整个计算过程建立在博弈论的理性参与者基础上，也就是

说，每个参与者都希望最大化自己的收益。由于与恶意攻击者采取攻击行为相关的阶段博弈包括Γ_V、Γ_W和Γ_C，因此，恶意攻击者的攻击行为预测就是要计算这些阶段博弈中恶意攻击者的混合策略纳什均衡，也就是说，要使得恶意攻击者在最大化自己的期望收益$E(\alpha^k,\beta^k)$基础上得到其在不同阶段博弈中的混合策略α^k，其中对

$$\forall \Gamma_k \in \Gamma,\quad E(\alpha^k,\beta^k) = \sum_{i\in A_1}\sum_{j\in A_2}\alpha_i^k\beta_j^k\mu_{ij}^k \tag{6-13}$$

设α^{k^*}和β^{k^*}分别为每个阶段博弈Γ_k中恶意攻击者和无线传感器网络系统采取的最优混合策略，因为本章定义的攻击预测随机博弈是零和的，所以这些最优混合策略可通过计算

$$\max_{\alpha^k}\min_{\beta^k} E(\alpha^k,\beta^k) \tag{6-14}$$

得到。这样，矩阵博弈Δ_k的值$\text{val}(\Delta_k)$就可以通过计算

$$\text{val}(\Delta_k) = E(\alpha^{k^*},\beta^{k^*}) \tag{6-15}$$

得到。最终的目标是要得到恶意攻击者在不同阶段博弈的整个最优混合策略集合

$$\alpha^* = \{\alpha^{V^*},\alpha^{W^*},\alpha^{C^*}\} \tag{6-16}$$

下面给出相应的基于攻击预测随机博弈的恶意攻击者攻击预测算法。

算法 6-1　计算恶意攻击者最优混合策略集合。

输入：$\mathbb{G}$

输出：α^*

1. 初始化期望收益向量$\boldsymbol{v}=(\nu_V,\nu_W,\nu_C)=(0,0,0)$。

2. 初使化矩阵博弈$\boldsymbol{\Delta}_C=(\nu_{ij}^C)=(\mu_{ij}^C)=(r_{ij}^C)$。

3. 由式(6-14)计算矩阵博弈$\boldsymbol{\Delta}_C$中恶意攻击者和无线传感器网络系统各自的最优混合策略α^{C^*}和β^{C^*}。

4. 设置$\nu_C=\text{val}(\boldsymbol{\Delta}_C)=E(\alpha^{C^*},\beta^{C^*})$。

5. 使用ν_C替换式(6-11)中的Γ_C并计算得到ν_{ij}^W，从而形成$\boldsymbol{\Delta}_W$。

6. 由式(6-14)计算矩阵博弈$\boldsymbol{\Delta}_W$中恶意攻击者和无线传感器网络系统各自的最优混合策略α^{W^*}和β^{W^*}。

7. 设置$\nu_W=\text{val}(\boldsymbol{\Delta}_W)=E(\alpha^{W^*},\beta^{W^*})$。

8. 使用ν_W替换式(6-10)中的Γ_W并计算得到ν_{ij}^V，从而形成$\boldsymbol{\Delta}_V$。

9. 由式(6-14)计算矩阵博弈$\boldsymbol{\Delta}_V$中恶意攻击者和无线传感器网络系统各自的最优混合策略α^{V^*}和β^{V^*}。

10. 返回$\alpha^*=\{\alpha^{V^*},\alpha^{W^*},\alpha^{C^*}\}$。

6.4　受攻击无线传感器网络的可生存性评估

6.4.1　基于连续时间马尔可夫链的传感器节点各状态转换关系

由于在无线传感器网络中，触发传感器节点转换状态的事件是随机的，这种特性使得连续时间马尔可夫链是建立传感器节点状态模型合适的工具。在受攻击无线传感器网络中，虽然恶意攻击者是否发动攻击的决策不是随机的，但发动攻击的时间和努力程度却是随机

分布的,况且在 6.3.2 小节已建立恶意攻击者的故意攻击与连续时间马尔可夫链中状态转换之间的联系,因此,将一个传感器节点的生命期看作是一个动态系统,从而可以应用连续时间马尔可夫链这种随机过程建立相应的模型。

在图 6-2 中,离散状态集合可表示为

$$S = \{H, V, W, C, F\} \tag{6-17}$$

设

$$X(t) = \{X_H(t), X_V(t), X_W(t), X_C(t), X_F(t)\} \tag{6-18}$$

式中,$X_i(t)$ 为在时刻 t 一个传感器节点处于状态 i 的概率,则描述一个传感器节点状态变化的表达式为

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = X(t)\boldsymbol{P} \tag{6-19}$$

式中,$\boldsymbol{P}$ 为不同状态之间的转换关系且是一个 5×5 的状态转换矩阵,其元素 p_{ij} 表示状态 i 和 j 之间的转换概率,可表示为

$$p_{ij} = \begin{cases} \lim\limits_{\mathrm{d}t\to 0}\left(\dfrac{\Pr(\text{chang from } i \text{ to } j \text{ in}(t, t+\mathrm{d}t))}{\mathrm{d}t}\right), & i \neq j \\ -\sum\limits_{j\neq i} p_{ij}, & i = j \end{cases} \tag{6-20}$$

于是,独立于初始状态的连续时间马尔可夫链稳态概率为

$$\boldsymbol{X} = \{X_H, X_V, X_W, X_C, X_F\} \tag{6-21}$$

可以从包含 5 个等式的方程组中解得,该方程组由

$$\boldsymbol{XP} = 0 \tag{6-22}$$

形成的 5 个等式中的任意 4 个等式再与第 5 个等式

$$\sum X_i = 1 \tag{6-23}$$

组合得到。

6.4.2 可靠度和生存期

无线传感器网络可靠度反映了在某个特定时间某个特定状态下传感器节点能持续数据感知、传输、融合等的概率。由于传感器节点一旦被破坏就很难修复,因此这里选择平均无故障时间这个指标来反映一个传感器节点的可靠度,再把平均无故障时间关联到一个传感器节点的故障率,就可以得到以串—并系统意义上的整个无线传感器网络的可靠度。

根据 Buzacott[312] 的方法,从一个连续时间马尔可夫链中通过计算获得平均无故障时间。将状态空间写成

$$S = \{S_{\text{Work}}, S_{\text{Failure}}\} \tag{6-24}$$

其中,

$$S_{\text{Work}} = \{H, V, W\} \tag{6-25}$$

表示传感器节点处于能够正常工作状态的集合。

$$S_{\text{Failure}} = \{C, F\} \tag{6-26}$$

表示传感器节点处于故障状态的集合。这样,矩阵 $\boldsymbol{P}$ 就可重写为

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{P}_1 & \boldsymbol{P}_2 \\ \boldsymbol{P}_3 & \boldsymbol{P}_4 \end{bmatrix} \tag{6-27}$$

式中，$\boldsymbol{P}_1$ 为由各个工作状态之间转换概率组成的 3×3 矩阵；$\boldsymbol{P}_2$ 为由各个工作状态转换到各个故障状态概率组成的 3×2 矩阵；$\boldsymbol{P}_3$ 为由各个故障状态转换到各个工作状态概率组成的 2×3 矩阵；$\boldsymbol{P}_4$ 为由各个故障状态之间转换概率组成的 2×2 矩阵。相应地，可将连续时间马尔可夫链的稳态概率改写成

$$\boldsymbol{X} = \{\boldsymbol{X}_{\text{Work}},\ \boldsymbol{X}_{\text{Failure}}\} \tag{6-28}$$

其中，

$$\boldsymbol{X}_{\text{Work}} = \{X_H,\ X_V,\ X_W\} \tag{6-29}$$

$$\boldsymbol{X}_{\text{Failure}} = \{X_C,\ X_F\} \tag{6-30}$$

给定时刻 $t=0$ 的任意一个工作状态 $\boldsymbol{X}_{\text{Work}}(0)$，一个传感器节点的平均无故障时间可以通过

$$\text{MTTF} = \boldsymbol{X}_{\text{Work}}(0)\,(-\boldsymbol{P}_1)^{-1}\boldsymbol{h} \tag{6-31}$$

计算得到，其中，

$$\boldsymbol{X}_{\text{Work}}(0) = \boldsymbol{X}_{\text{Work}}/\boldsymbol{X}_{\text{Work}}\boldsymbol{h} \tag{6-32}$$

$$\boldsymbol{h} = [1 \quad 1 \quad 1]^{-1} \tag{6-33}$$

给定一个复杂系统中的单个组件 i，其可靠度 $R_i(t)$ 和平均无故障时间都可以从它的故障率 $\lambda_i(t)$ 计算得到。为简化起见，假设在系统运行的整个时期内故障率是一个常量，即

$$\lambda_i(t) = \lambda \tag{6-34}$$

则可得

$$\lambda = \frac{1}{\text{MTTF}} \tag{6-35}$$

$$R_i(t) = \exp(-\lambda t) \tag{6-36}$$

这里将无线传感器网络中的一个传感器节点看作一个复杂系统中的单个组件，并由于这些传感器节点具有类似的特性，所以可假设它们具有相同的故障率 λ，并假设每个簇包含 m 个传感器节点。根据 6.3 节的分析，整个无线传感器网络可以看作由多个簇组成的串—并系统，显然，每个传感器节点出现故障的概率是相互独立的，一个簇内只有所有的候选簇头都出现故障才会使该簇失去正常工作的能力。因此，可得到一个簇的可靠度为

$$R_{C_i}(t) = 1 - \prod_{i=1}^{m}(1 - R_i(t)) = 1 - (1 - \exp(-\lambda t))^m \tag{6-37}$$

因为一条路由上任意一个簇出现故障就会导致整条路由失败，所以可得到一条路由的可靠度，即

$$R_{R_i}(t) = \prod_{i=1}^{n} R_{C_i}(t) = (1 - (1 - \exp(-\lambda t))^m)^n \tag{6-38}$$

式中，n 为一条路由经过的簇数。假设任意一条传输数据的路由所经过的簇数相同，则整个无线传感器网络的可靠度为

$$R(t) = 1 - \prod_{i=1}^{l}(1 - R_{R_i}(t)) = 1 - ((1 - (1 - \exp(-\lambda t))^m)^n)^l \tag{6-39}$$

式中，l 为整个无线传感器网络中从源节点到基站的所有可用路由数。相应的整个无线传感器网络的生存期为

$$ST = \int_0^{\infty} R(t)\,\mathrm{d}t = \int_0^{\infty} (1 - ((1 - (1 - \exp(-\lambda t))^m)^n)^l)\,\mathrm{d}t \tag{6-40}$$

6.4.3 稳态可用度

虽然整个无线传感器网络的可用度与可靠度密切相关，但它代表了传感器节点在给定时间间隔中任意的一个时刻具备执行指定功能如数据感知、传输、融合等的能力。实际上，可靠度代表了一段时间内处于正常工作的能力，而可用度代表了当任务在任意一个时刻需要启动时处于正常工作的能力。这里选择连续时间马尔可夫链为传感器节点的整个生命期建立模型，因此主要考虑整个无线传感器网络的稳态可用度。从数学语言的角度来讲，稳态可用度代表了当时间无限扩展时瞬时可用度函数的极限值，即

$$A_{S_i} = \lim_{t \to \infty} A_{I_i}(t) \tag{6-41}$$

式中，A_{S_i}为传感器节点 i 的稳态可用度；瞬时可用度函数 $A_{I_i}(t)$表示传感器节点 i 在时刻 t 能正常提供所需功能的概率。在稳态可用度的实际计算过程中，可通过计算连续时间马尔可夫链的稳定点得到，此时一个传感器节点的可用度是一个常量值，即

$$A_{S_i} = X_H + X_V + X_W \tag{6-42}$$

与 6.4.2 小节类似的推导可得到整个无线传感器网络的稳态可用度为

$$Av = 1 - ((1 - (1 - A_{S_i})^m)^n)^l \tag{6-43}$$

6.5 实验

使用 MATLAB R2010a，接下来分析不同的博弈参数如何影响恶意攻击者的期望动机，以及一个受攻击传感器节点的平均无故障时间在状态 V 和 W 如何依赖于恶意攻击者的期望动机。最后，从可靠度、生存期、稳态可用度 3 个方面给出整个无线传感器网络可生存性评估。

6.5.1 恶意攻击者的期望动机

无线传感器网络攻击预测随机博弈中的不同动作策略将影响恶意攻击者的期望动机。当恶意攻击者选择动作 a 而无线传感器网络系统选择动作 ϕ 时，恶意攻击者将得到一个正收益，恶意攻击者在接下来的攻击继续成功的话，他将进一步获得更多的正收益。另外，当恶意攻击者选择动作 a 而无线传感器网络系统选择动作 d 时，恶意攻击者将遭受损失。因此最终收益值会影响恶意攻击者在选择动作 a 还是 ϕ 的期望动机的决策。

根据算法 6-1，恶意攻击者的期望动机 $\alpha^* = \{\alpha^{V^*}, \alpha^{W^*}, \alpha^{C^*}\}$是从阶段博弈 Γ_V、Γ_W、Γ_C 中计算得到的最优混合策略组成的集合。虽然在不同的阶段博弈中存在不同的收益值，但只要选择任意一个阶段博弈 Γ_k，就可说明博弈参数将如何影响恶意攻击者的期望动机，为说明这种影响，假设任意的一个阶段博弈 Γ_k 的支付矩阵如表 6-1 所示。

表 6-1 阶段博弈 Γ_k 的支付矩阵

动作	Non-defend(ϕ)	Defend(d)
Attack (a)	1	r_{ad}
Non-attack (ϕ)	$r_{\phi\phi}$	0

在区间[−10,0]上分别改变表 6-1 中 r_{ad} 和 $r_{\phi\phi}$ 的值，就可以得到恶意攻击者的期望动机是如何由“动作对”(a, d)和(ϕ,ϕ)带来的负收益决定的。在决策过程中，可利用式(6-14)求解 Γ_k 的纳什均衡得到恶意攻击者的最优混合策略 $\alpha^{k^*}=(\alpha_a^{k^*},\alpha_\phi^{k^*})$。图 6-3 给出了恶意攻击者的最优攻击概率 $\alpha_a^{k^*}$ 与博弈参数 r_{ad} 和 $r_{\phi\phi}$ 之间的关系。

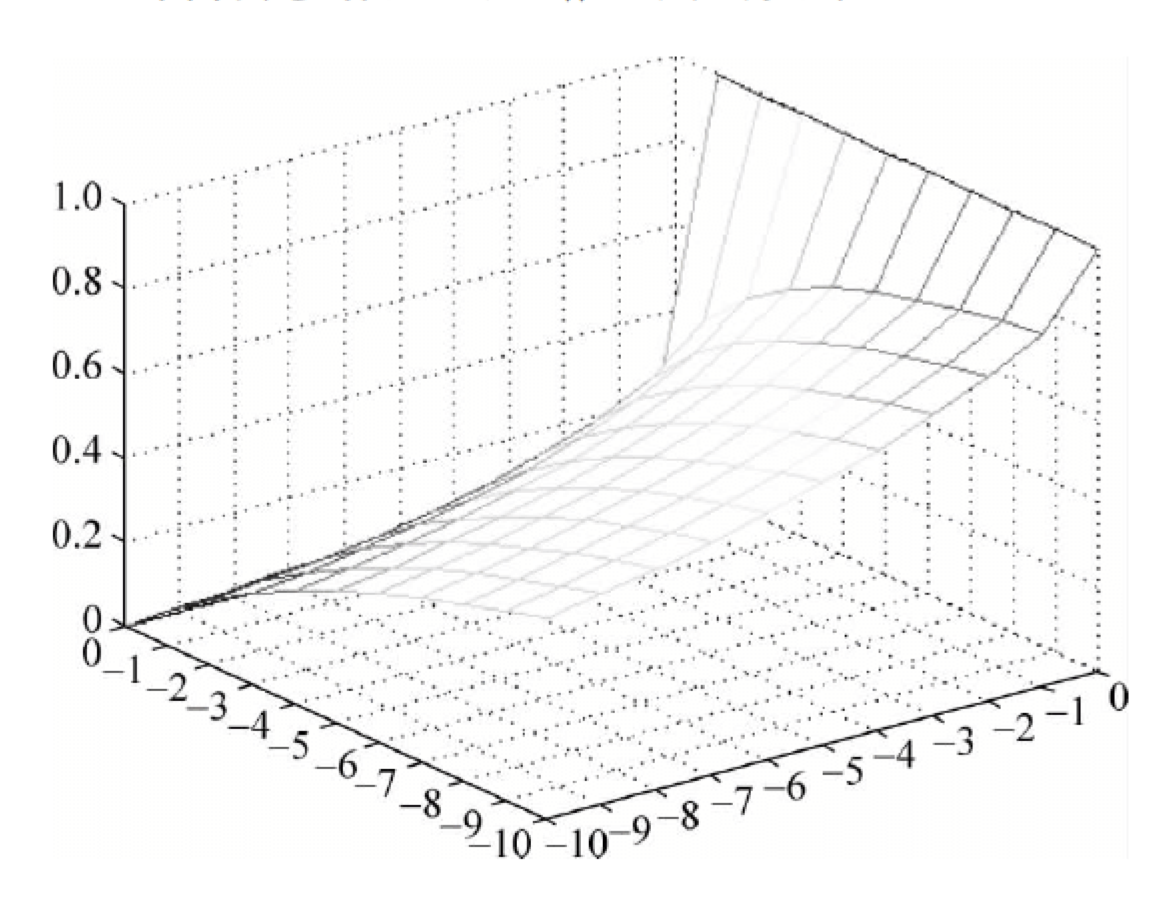


图 6-3　$\alpha_a^{k^*}$ 与博弈参数 r_{ad} 和 $r_{\phi\phi}$ 之间的关系

在图 6-3 中，x 轴、y 轴和 z 轴分别表示 r_{ad}、$r_{\phi\phi}$ 和 $\alpha_a^{k^*}$。首先，设 $r_{\phi\phi}$ 是静态的而 r_{ad} 是动态变化的，就可以得到最优攻击概率 $\alpha_a^{k^*}$ 的相应变化。例如，当 $r_{ad}=-1$ 且 $r_{\phi\phi}=-5$ 时，$\alpha_a^{k^*}\approx 0.7143$；当 $r_{ad}=-6$ 且 $r_{\phi\phi}=-5$ 时，$\alpha_a^{k^*}\approx 0.4167$。这表明最优攻击概率随着“动作对”$(a, d)$损失的增大而减小。其次，设 r_{ad} 是静态的而 $r_{\phi\phi}$ 是动态变化的，可以看到最优攻击概率随着“动作对”(ϕ,ϕ)损失的增大而增大。例如，当 $r_{ad}=-1$ 且 $r_{\phi\phi}=-1$ 时，$\alpha_a^{k^*}\approx 0.3333$；当 $r_{ad}=-1$ 且 $r_{\phi\phi}=-8$ 时，$\alpha_a^{k^*}\approx 0.8000$。最后来观察 $\alpha_a^{k^*}$ 何时达到极值。显然，当 $r_{ad}=0$ 时，$\alpha_a^{k^*}=1$，这意味着恶意攻击者的最优攻击概率达到最大。实际上，此时动作 a 是严格占优的，因此恶意攻击者将始终选择动作 a。另外，当 $r_{\phi\phi}=0$ 且 $r_{ad}<0$ 时，$\alpha_a^{k^*}=0$，这意味着恶意攻击者的最优攻击概率达到最小，此时恶意攻击者无意选择动作 a，因为与选择动作 ϕ 导致的零收益相比，选择动作 a 将获得负的期望收益值。

6.5.2　受攻击传感器节点的平均无故障时间

根据式(6-31)，一个受攻击传感器节点的平均无故障时间的影响因素主要是连续时间马尔可夫链达到稳态时的各状态概率以及处于正常工作状态的状态转移矩阵，这些概率可通过计算式(6-22)和式(6-23)得到。另外，状态转移矩阵 $\boldsymbol{P}$ 中的状态转移概率 p_{VW} 和 p_{WC} 分别与最优攻击概率 $\alpha_a^{V^*}$ 和 $\alpha_a^{W^*}$ 有关，$p_{ij}(i=j)$可以从式(6-20)获得。因此，接下来，根据 $\alpha_a^{V^*}$ 和 $\alpha_a^{W^*}$ 的变化讨论对受攻击传感器节点平均无故障时间的影响，结果如图 6-4 所示。

在图 6-4 中，x 轴、y 轴和 z 轴分别表示 $\alpha_a^{V^*}$、$\alpha_a^{W^*}$ 和一个受攻击传感器节点的平均无故障时间。根据经验值，假设 $\eta=1/300$，$\theta_{VW}=1/3$，$\theta_{WC}=3$，$p_{WV}=1/50$，$p_{VH}=1/3$，$p_{WH}=1/40$，$p_{CH}=1/100$ 和 $p_{FH}=1/200$（每时）。由图 6-2，$p_{HW}=p_{HC}=p_{VC}=p_{WV}=p_{CV}=p_{CW}=p_{FV}=$

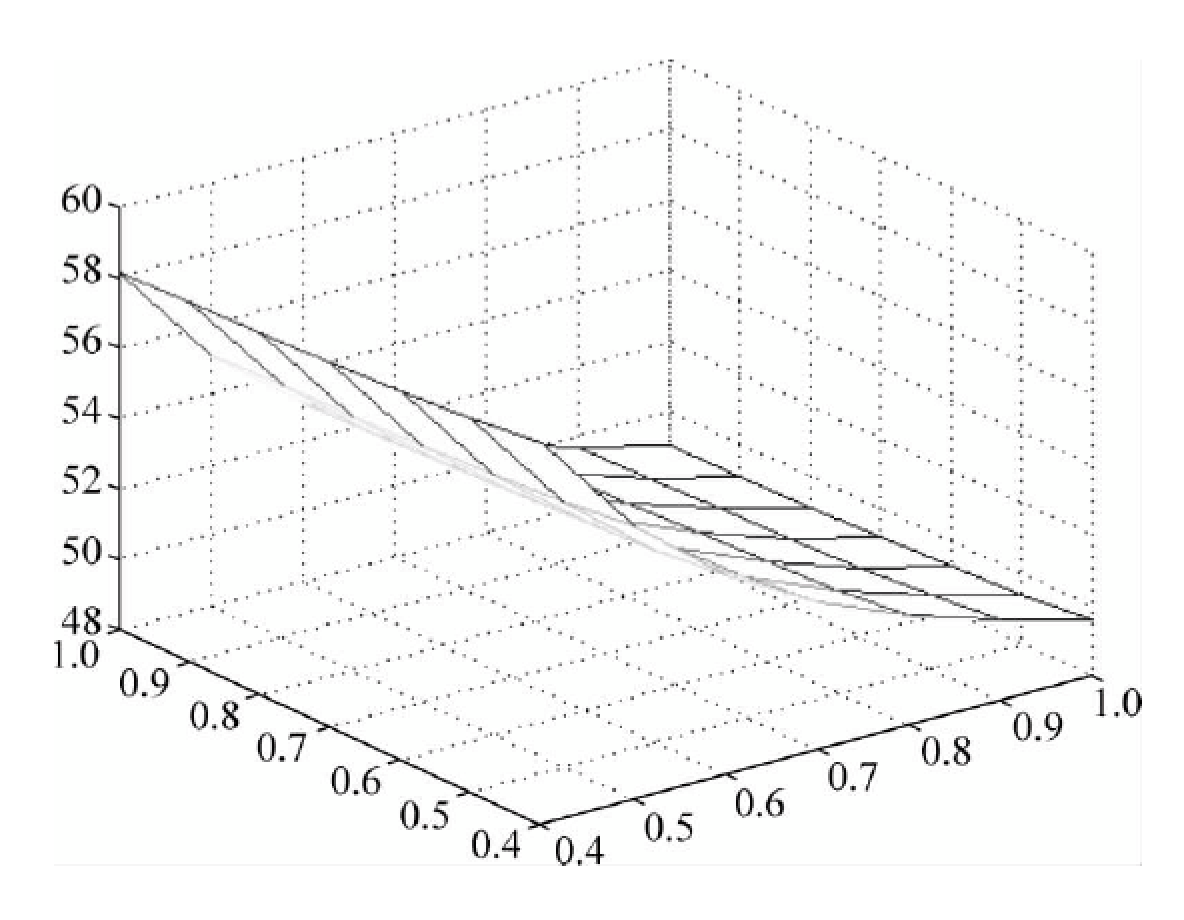


图 6-4 平均无故障时间与 $\alpha_a^{V^*}$ 和 $\alpha_a^{W^*}$ 的关系

$p_{FW}=p_{FC}=0$。因此，状态转移矩阵为

$$\boldsymbol{P}=\begin{bmatrix} -(p_{HV}+\eta) & p_{HV} & 0 & 0 & \eta \\ p_{VH} & -(p_{VH}+\alpha_a^{V^*}\theta_{VW}+\eta) & \alpha_a^{V^*}\theta_{VW} & 0 & \eta \\ p_{WH} & 0 & -(p_{WH}+\alpha_a^{W^*}\theta_{WC}+\eta) & \alpha_a^{W^*}\theta_{WC} & \eta \\ p_{CH} & 0 & 0 & -(p_{CH}+\eta) & \eta \\ p_{FH} & 0 & 0 & 0 & -p_{FH} \end{bmatrix},$$

其中，最优攻击概率 $\alpha_a^{V^*}$ 和 $\alpha_a^{W^*}$ 由算法 6-1 计算得到。为了简化起见，假设 $\alpha_a^{V^*}$ 和 $\alpha_a^{W^*}$ 在区间[0.4，1]之间变化。

从图 6-4 中可看出，当恶意攻击者在状态 V 选择动作 a 的概率持续增加时，一个受攻击传感器节点的平均无故障时间显著减小。例如，当 $\alpha_a^{W^*}=0.4$ 且 $\alpha_a^{V^*}=0.4$ 时，MTTF≈58.845h；当 $\alpha_a^{W^*}=0.4$ 且 $\alpha_a^{V^*}=1$ 时，MTTF≈49.585h，与前者相比将近减少了 15.74%。另外，当恶意攻击者在状态 W 改变攻击概率时，对一个受攻击传感器节点的平均无故障时间的影响微乎其微。例如，在 $\alpha_a^{V^*}=0.4$ 前提下，让 $\alpha_a^{W^*}$ 从 0.4 变化到 1，MTTF 的值仅细小地从 58.845h 变化到 58.182h。这些实验结果反映出在状态 V 恶意攻击者期望动机对一个受攻击传感器节点的平均无故障时间影响要大，因此在状态 V 预测恶意攻击者的期望动机比在状态 W 更重要。

6.5.3 整个无线传感器网络的可靠度和生存期

根据式(6-39)，整个无线传感器网络的可靠度与一个受攻击传感器节点的故障率、一个簇内的传感器节点数、一条路由上所经过的簇头数和整个无线传感器网络可用的路由数等都有关联。故障率 λ 可从平均无故障时间的表达式(6-35)计算得到，接下来将讨论 $R(t)$ 和不同参数 m、n 和 l 值的关系。为得到一个受攻击传感器节点的平均无故障时间，首先应该计算 $\alpha_a^{V^*}$ 和 $\alpha_a^{W^*}$。根据 6.5.2 小节中给定的经验值，可得到

$$q_{a\phi}^{VW}=\alpha_a^V\theta_{VW}/(\alpha_a^V\theta_{VW}+\rho_{VH}+\eta)\approx 0.8621 \tag{6-44}$$

$$q_{a\phi}^{WC}=\alpha_a^W\theta_{WC}/(\alpha_a^W\theta_{WC}+\rho_{WH}+\eta)\approx 0.9906 \tag{6-45}$$

设 $r_{a\phi}^{V}=r_{a\phi}^{W}=r_{a\phi}^{C}=1$，$r_{ad}^{V}=-4$，$r_{ad}^{W}=-3$，$r_{ad}^{C}=-2$，$r_{\phi\phi}^{V}=r_{\phi\phi}^{W}=r_{\phi\phi}^{C}=-5$ 和 $r_{\phi d}^{V}=r_{\phi d}^{W}=r_{\phi d}^{C}=0$，相应地阶段博弈就可分别写成

$$\Gamma_V=\begin{pmatrix}1+0.8621\Gamma_W & -4\\ -5 & 0\end{pmatrix} \tag{6-46}$$

$$\Gamma_W=\begin{pmatrix}1+0.9906\Gamma_C & -3\\ -5 & 0\end{pmatrix} \tag{6-47}$$

$$\Gamma_C=\begin{pmatrix}1 & -2\\ -5 & 0\end{pmatrix} \tag{6-48}$$

通过算法 6-1 可解攻击预测随机博弈，得到

$$\alpha_a^{*}=\{\alpha_a^{V^*},\alpha_a^{W^*},\alpha_a^{C^*}\}\approx\{0.6000,\ 0.6442,\ 0.6250\} \tag{6-49}$$

这些在不同阶段博弈的最优攻击概率被用于状态转移矩阵 $\boldsymbol{P}$ 中后，就可逐步计算得到整个无线传感器网络的可靠度和生存期。

实验结果如图 6-5 和图 6-6 所示，其中图 6-5 给出了当 $n=3$ 且 $l=3$ 时，一个簇内传感器节点分别为 2、4、6 对应的整个无线传感器网络可靠度变化曲线；图 6-6 给出了当 $n=6$ 且 $l=3$ 时，一个簇内传感器节点数分别为 2、4、6 对应的整个无线传感器网络可靠度变化曲线。

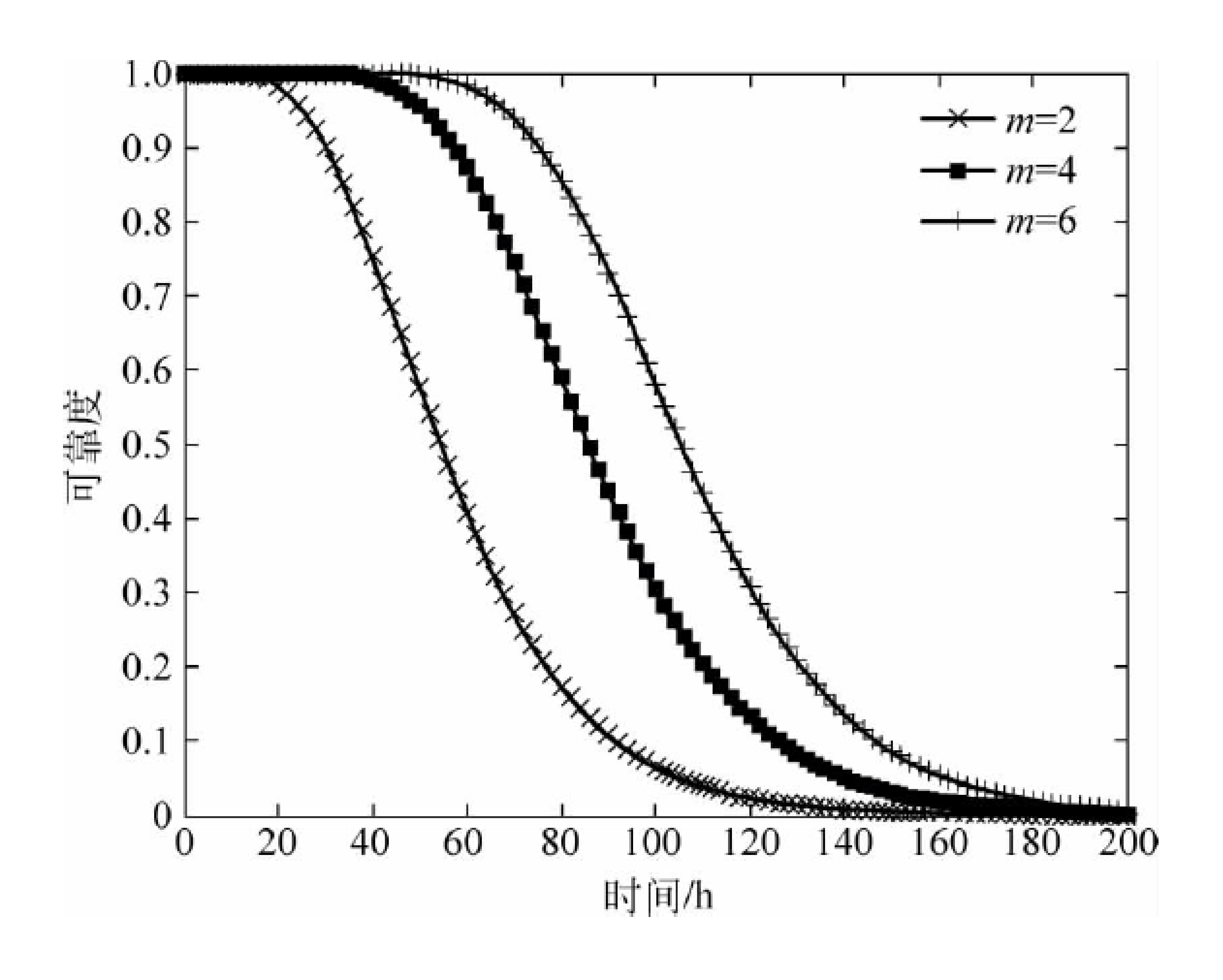


图 6-5　当 $n=3$ 且 $l=3$ 时整个无线传感器网络的可靠度

从图 6-5 可以看出，当 n 和 l 为静态值时，一个簇内的传感器节点数越多，整个无线传感器网络的可靠度越大。例如，在恶意攻击者的攻击环境中，要使整个无线传感器网络的可靠度降到 0.5，当一个簇内的传感器节点数分别为 2、4、6 时，需要的时间分别为 55h、83h 和 107h。与图 6-6 比较，当一条路由所经过的簇数从 3 增加到 6 时，整个无线传感器网络的可靠度明显下降。例如，当 $m=2$ 时，要使整个无线传感器网络的可靠度下降到 0.5，在图 6-5 所示的环境中需要 55h，而在图 6-6 所示的环境中只需要 35h，大约减少了 36.36%。这些结果反映出为了提高整个无线传感器网络的可靠度，应该增加同一个簇内的传感器节点数而尽量减少一条路由需经过的簇数。

下面讨论整个无线传感器网络生存期的变化情况，图 6-7 至图 6-9 分别给出了 l 为静态值且 $l=3$、m 为静态值且 $m=3$ 和 n 为静态值且 $n=3$ 对应的生存期变化曲线。

在图 6-7 中，整个无线传感器网络生存期随着一个簇内传感器节点数的增加和一条路由上经过簇数的减少而增加，但这些变化趋势也有很大的不同。例如，当 $m=3$ 时，随着 n 值从 8 变化 7，整个无线传感器网络的生存期从 45h 缓慢增长到 48h；而随着 n 值从 3 变化到 2，整个无线传感器网络的生存期从 76h 迅速增加到 96h。又如，当 $n=3$ 时，随着 m 值从 2 变化到 3，整个无线传感器网络的生存期迅速从 58h 增加到 76h；而随着 m 值从 6 变化到 7，整个无线传感器网络的生存期仅缓慢地从 109h 增加到 116h。在图 6-8 中，可以看到整个无线传感器网络的生存期随着一条路由经过的簇数的减少和整个无线传感器网络路由数的减少而增加。而在图 6-9 中，可以看到整个无线传感器网络的生存期随着一个簇内传感器节点数的增加和整个传感器网络路由数的增加而增加。

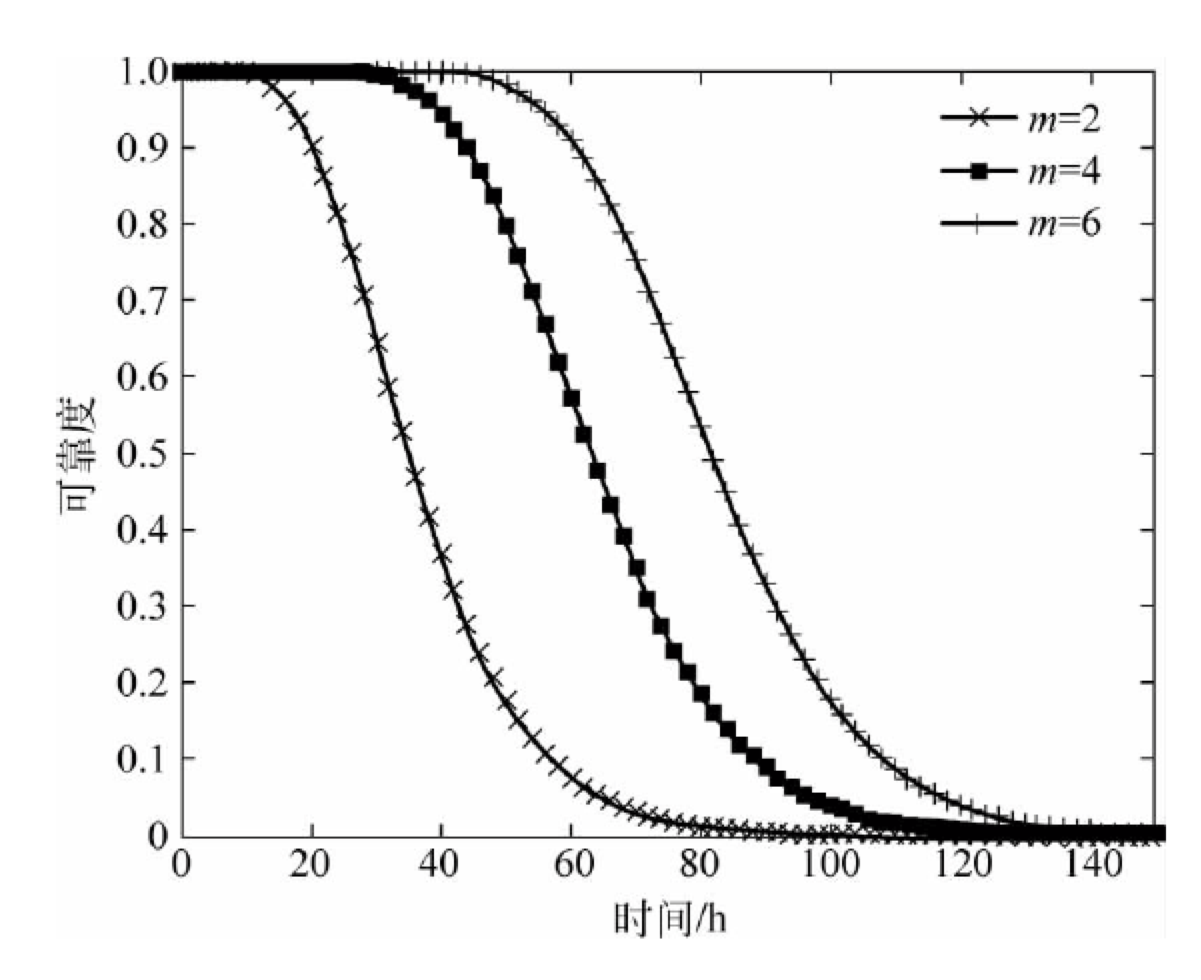


图 6-6 当 $n=6$ 且 $l=3$ 时整个无线传感器网络的可靠度

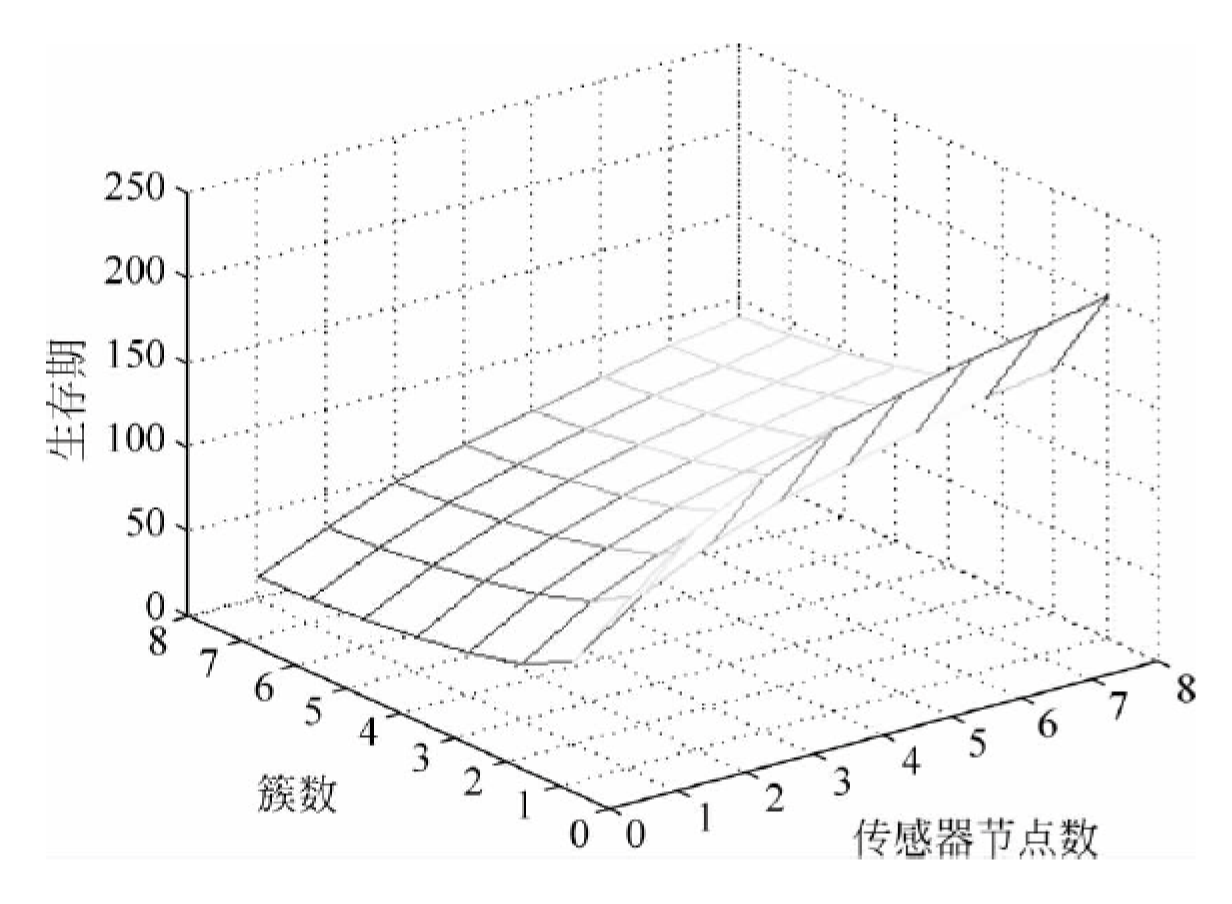


图 6-7 当 $l=3$ 时的整个无线传感器网络的生存期

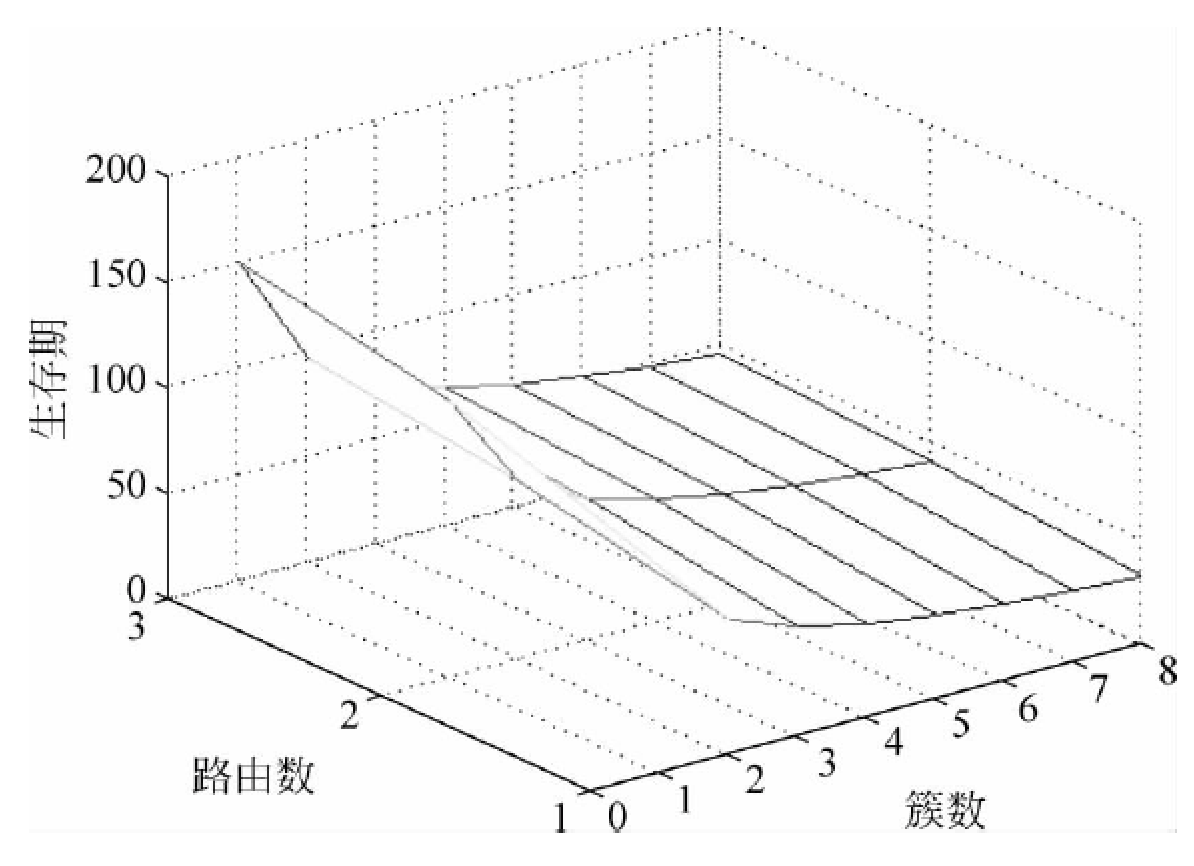


图 6-8　当 $m=3$ 时的整个无线传感器网络的生存期

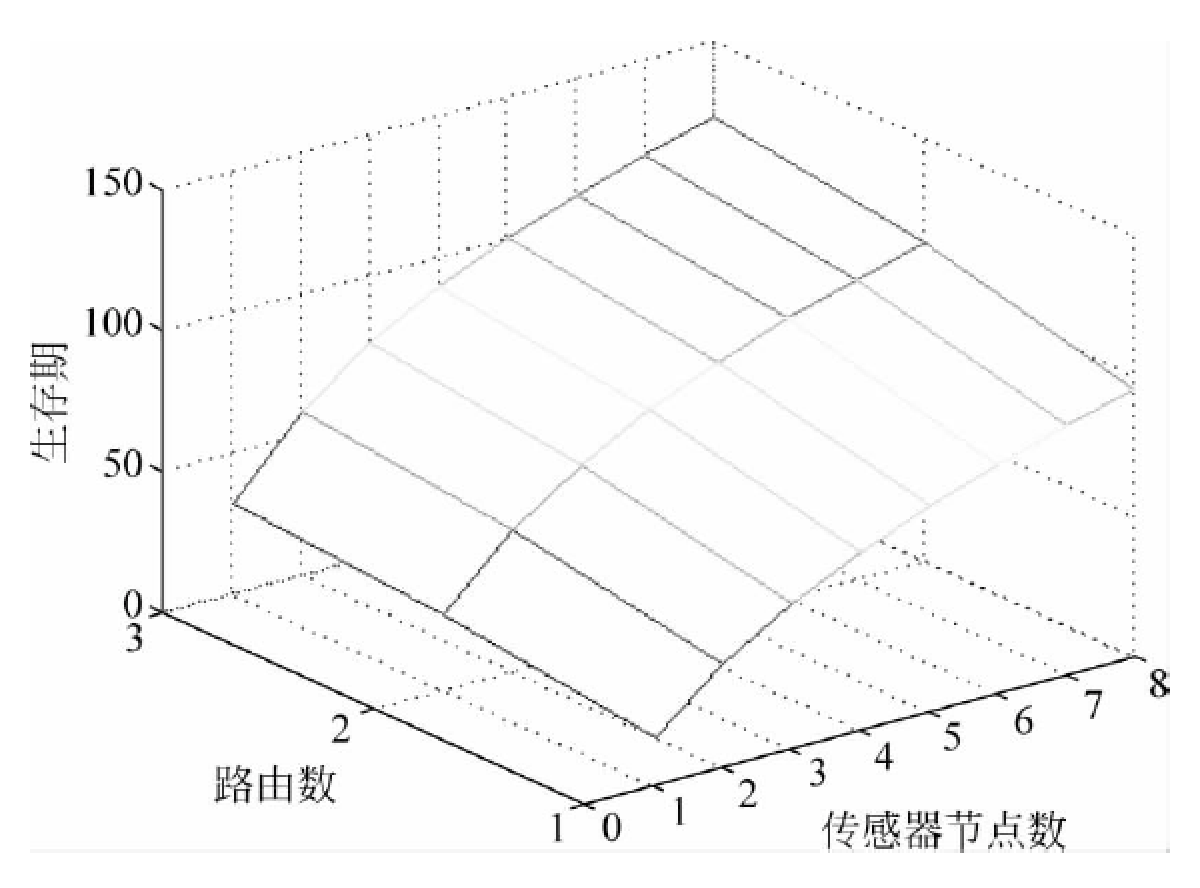


图 6-9　当 $n=3$ 时的整个无线传感器网络的生存期

6.5.4　稳态可用度

与 6.5.3 小节整个无线传感器网络生存期实验类似，整个无线传感器网络的稳态可用度实验也包括 3 个方面，结果如图 6-10～图 6-12 所示。

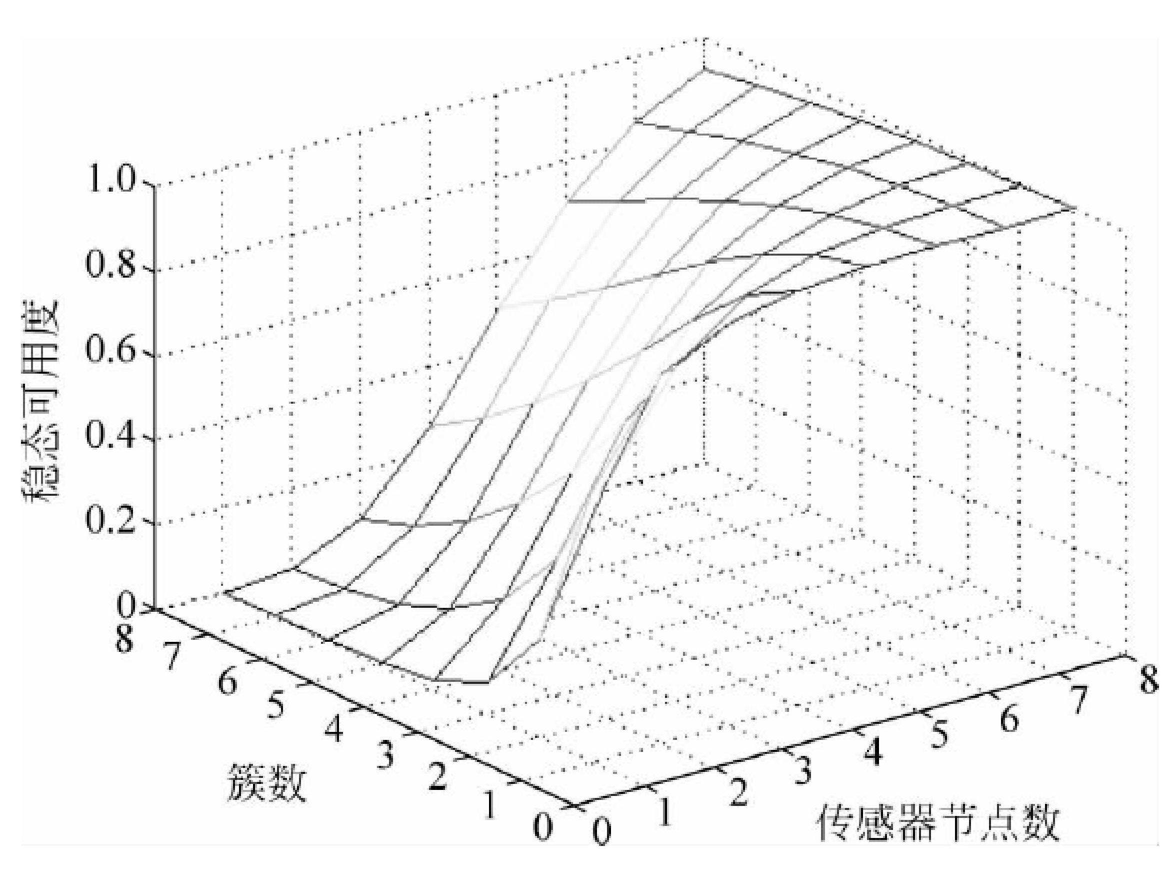


图 6-10　当 $l=3$ 时整个无线传感器网络的稳态可用度

在图 6-10 中，整个无线传感器网络的稳态可用度随着一个簇内传感器节点数的增加和一条路由经过簇数的减小而增大。与整个无线传感器网络的生存期变化趋势类似的是，稳态可用度的变化趋势也有很大的不同。例如，当 $m=3$ 时，随着 n 值从 8 变化到 7，整个无线传感器网络的稳态可用度缓慢地从 0.0832 增加到 0.1277；而随着 n 值从 3 变化到 2，整个无线传感器网络的稳态可用度迅速地从 0.6005 增加到 0.7957。又例如，当 $n=3$ 时，随着 m 值从 2 变化到 3，整个无线传感器网络的可用度迅速地从 0.3215 增加到 0.6005；而随着 m 值从 6 变化到 7，整个无线传感器网络的稳态可用度微小地从 0.9611 增加到 0.9843。在图 6-11 中，整个无线传感器网络的稳态可用度随着一条路由经过的簇数的减少和整个无线传感器网络路由数的减少而增加。而在图 6-12 中，整个无线传感器网络的稳态可用度随着一个簇内传感器节点的增加和整个无线传感器网络路由数的增加而增加。另外，还可以看出，一个簇内传感器节点数和一条路由经过的簇数对整个无线传感器网络稳态可用度的影响比整个无线传感器网络的路由数要大。

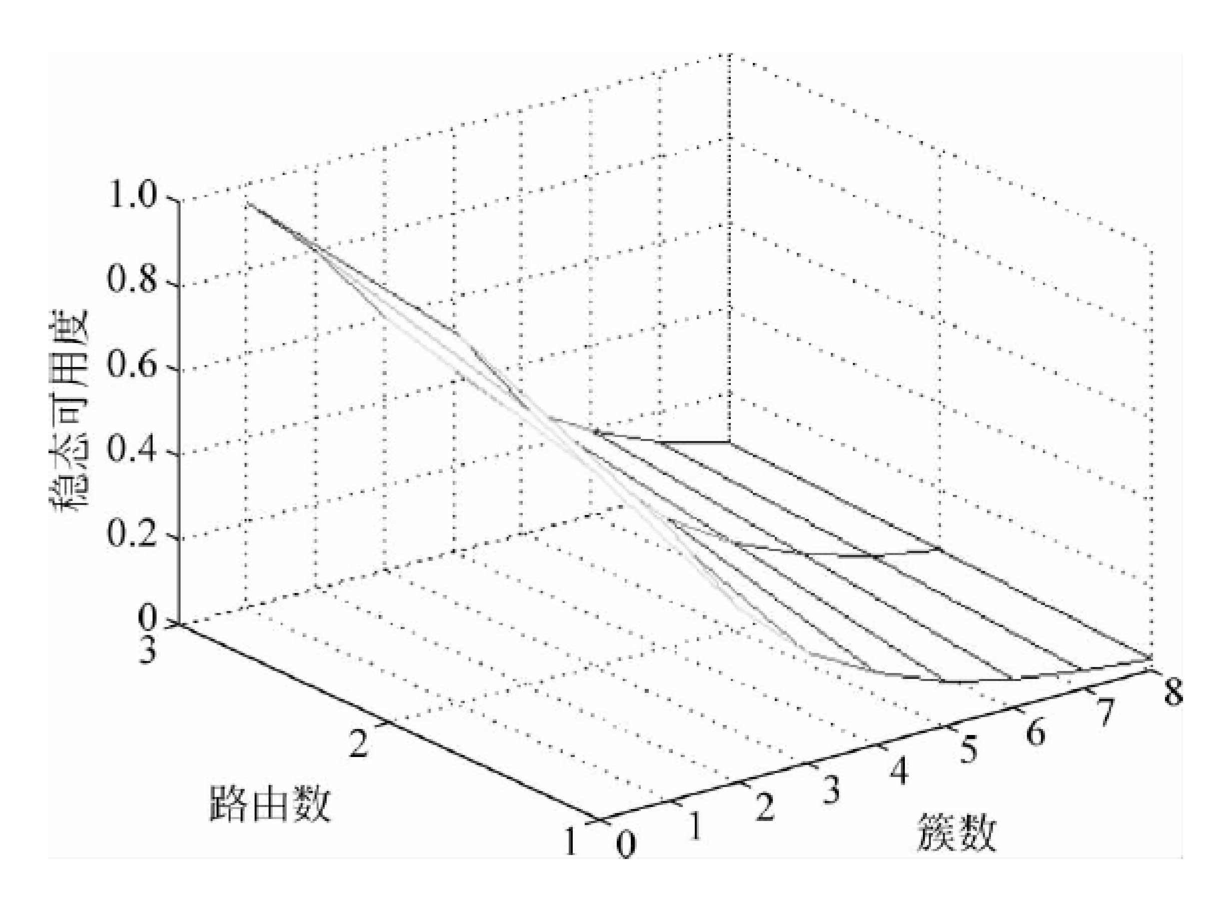


图 6-11 当 $m=3$ 时整个无线传感器网络的稳态可用度

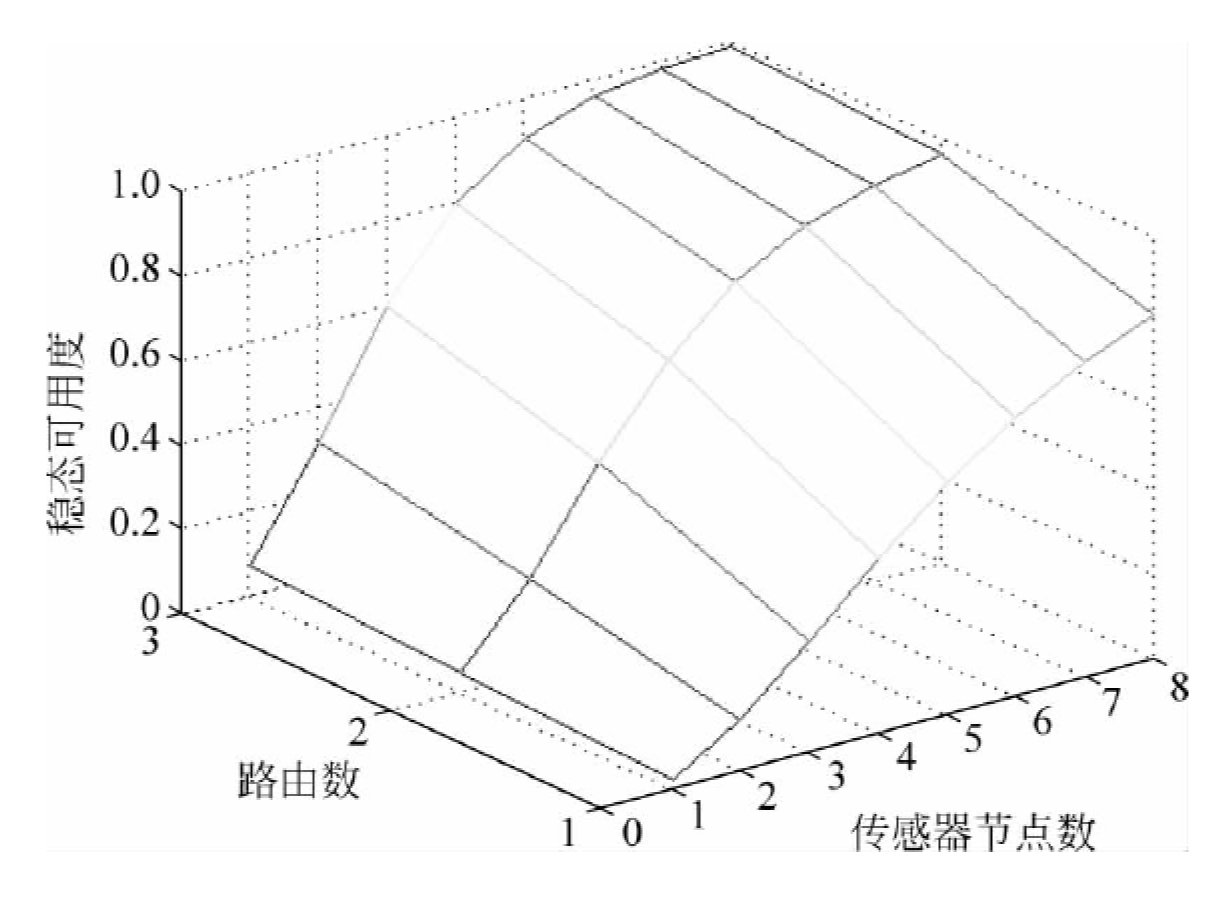


图 6-12 当 $n=3$ 时整个无线传感器网络的稳态可用度

6.6 小结

为构建高可生存的无线传感器网络，本章提出了一种面向受攻击无线传感器网络的可生存性评估机制。本章选择能量效率较高的聚簇无线传感器网络作为研究主体，并把它看作可靠性理论中的一个串—并系统，通过提出一种能得到理性恶意攻击者期望动机的攻击预测随机博弈，解决了恶意攻击者攻击行为的故意性与连续时间马尔可夫链中的随机性之间的矛盾。得到的恶意攻击者最优攻击概率被成功用于连续时间马尔可夫链中的状态转移矩阵，从而成功地建立了受攻击传感器节点的生命期模型。在此基础上，依次得到了一个受攻击传感器节点的平均无故障时间、可靠度、生存期和稳态可用度，这些评估指标反映了受攻击无线传感器网络的可生存特性。实验结果表明，恶意攻击者的期望动机与攻击预测随机博弈中的博弈参数有很大的联系，并且得到了为有效延长受攻击传感器节点的平均无故障时间需要着重在状态 V 而不是状态 W 进行预测的结论。同时，也验证了本章提出的可生存性评估机制对受攻击无线传感器网络可生存性评估的有效性，从而为设计高可生存的无线传感器网络提供了理论基础。

第 7 章

无线传感器网络受攻击协调器节点的防御响应博弈机制研究

在 ZigBee 无线传感器网络中，协调器节点是控制整个网络运行的关键节点，它决定着网络的协作通信性能。恶意节点为了耗费网络资源和干扰网络运行，选择协调器节点作为攻击目标，一旦攻击成功将导致多条通信路径的源和目的节点之间的一轮通信失败。协调器节点选择技术不仅要考虑数据包分发延迟、网络生命期等因素，而且还要考虑网络攻击防御策略。在本章中，为了最小化从源到目的节点的数据包分发平均跳数并且延长网络生命期，提出了基于博弈论和模糊逻辑的协调器节点选择算法。在此算法中，首先使用随机博弈对攻击进行动态响应，然后通过模糊逻辑选择通信质量较好的传感器节点作为协调器节点，最终提高网络的服务质量和安全性。

7.1 引言

在环境监控[355]、建筑物监控[356]和智能家庭[357]等应用领域，ZigBee 无线传感器网络的安全性和服务质量（QoS）影响着其在各种应用中的数据分发性能。传感数据的传输总是面临许多的恶意攻击[358]，为了使其更加安全，已提出了多种安全协议解决对应的网络安全问题[359]，这些协议包括 802.1X 端口访问控制协议、IP 安全协议等[360]。显然，对于不同种类的网络服务应用，保证 QoS 始终是一个关键因素，它包括传输延迟、网络吞吐量和数据丢包率等。

ZigBee 无线传感器网络由协调器节点和终端节点组成。网络拓扑一般为标准的簇树，树中的节点与协调器节点相连，从树根协调器节点到叶节点构成父子节点关系，能够彼此协作并提供有效和高效的数据分发服务。为了获得较高的数据分发性能，传感服务节点对 QoS 和安全有一定的需求。在不同的环境和时间点，要使有限的网络资源同时满足安全性和 QoS 需求是不现实的。对于提高安全性的组合服务，仅仅考虑传输成本而不考虑协调器节点的攻击也是不够的。为了延长网络生命期，避免由于攻击而造成的单点失败，需以分布的方式，基于协调器节点选择的方法进行服务组合。其中协调器节点将主动接收传感器节点的数据，然而，当协调器节点不在可信状态时，数据接收率将变小，QoS 将变低。因此，为了提高 QoS，可使用多个传感器节点共享可靠协调器节点的方法，使得传感数据的传输成本减少，即多个传感器节点形成一个共享协调器节点的联盟。当同一联盟的协调器节点未

受攻击且处于可靠状态时，每个传感器节点能够访问协调器节点；反之，协调器节点状态由正常转换为失败，变得无法访问。此时，为了可以在联盟中正常地分发数据，需使用处理协调器节点失败的容错机制来选择新的协调器节点。在选择时应考虑以下两方面问题：

（1）为了最小化能量成本，如何形成一个被联盟可靠共享的协调器节点？

（2）在满足 QoS 需求的前提下如何防御攻击者？

为了解决这两个问题，本章基于博弈论提出了协调器节点选择机制，实现针对恶意节点攻击的动态防御。其中，随机博弈被用于对攻击进行动态响应，演化博弈被用于选择协调器节点。

在扩展作者前期工作[361]的基础上，本章的工作主要包括以下内容：

（1）针对协调器节点的攻击，防御者能主动选择可靠的协调器节点来最小化网络能量损失，实现主动防御机制。

（2）将对网络协调器节点的攻击防御问题形式化为一个 2-player 零和博弈，其中的收益为 ZigBee 无线传感器网络的网络效用。

（3）运用演化博弈分析防御策略的响应过程，其中，网络中的节点作为一个博弈参与者，来自于不同邻居节点的局部组合估计信息作为协调器节点选择的依据。

（4）使用模糊逻辑为协调器节点选择提出了一个新的状态估计算法。使用演化博弈和随机博弈为协调器节点的攻击防御得到了相应的混合策略解，从而实现最大化博弈参与者收益的目的。

本章其余章节安排如下：7.2 节介绍相关工作；7.3 节描述系统模型；7.4 节提出基于演化博弈和随机博弈的动态攻击响应和协调器节点选择策略；7.5 实现数值仿真；7.6 节给出本章小结。

本章涉及的符号含义如下：

S 表示由协调器节点管理的传感器节点组合。

$s[i]$表示联盟传感器节点。

av_i 表示联盟传感器节点 $s[i]$的可用性。

av_M 表示整个联盟 M 的可用性。

p_M 表示使用每个传感器节点时产生的费用。

$c_M(s)$表示联盟 M 的成本。

$c_{s[i]}$表示节点 $s[i]$加入联盟的成本。

h_i 表示从联盟节点 $s[i]$到协调器节点的平均跳数。

E_r表示接收单个数据包的能量成本。

E_t表示发送单个数据包的能量成本。

$v_M(s)$表示联盟 M 的传感器节点组合成本。

$\widetilde{G}$表示协调器节点攻击响应随机博弈模型。

N 表示博弈参与者集合。

Z 表示状态空间。

$\{A_k|k\in N\}$表示参与者 k 采取的行动集合。

a_1 表示协调器节点状态从 NormalState 到 HackedState 状态的攻击行动。

r_1 表示攻击者在 NormalState 状态实施的攻击行动。

d_2 表示攻击行动 a_1 被防御者 defender 成功检测。

$\varnothing_2$ 表示攻击行动未被防御者 defender 成功检测。

$\{u_k | k \in N\}$表示参与者 k 的效用函数。

$u_2(a_2, a_{-2})$表示防御行动所获得的期望收益。

a_2 表示防御者 defender 采取的防御行动。

a_{-2}表示攻击者 attacker 采取的攻击行动。

A_1 表示攻击者 attacker 的行动集合。

A_2 表示防御者 defender 的行动集合。

$p(a)$表示防御者 defender 在集合 A_2 中采取防御行动的概率。

$u_2(a)$表示防御者 defender 在集合 A_2 中采取防御行动后，成功检测攻击获得的收益。

$\tau_i(x_i)$表示联盟节点 $s[i]$可用资源的损失率。

r_i 表示联盟节点 $s[i]$防御攻击的容侵性。

x_i、y_i 分别表示分配给防御者 defender 和攻击者 attacker 的资源。

m 表示联盟节点 $s[i]$可用资源的线性和非线性损失率。

α_i 表示防御者 defender 保护协调节点的困难程度。

$\tau_M(s)$表示联盟的可用资源的损失率。

$v'_M(s)$表示在考虑联盟可用资源损失率之后的联盟 M 的传感器节点组合成本。

d_i 表示协调器节点的度。

$h_{i,t+1}$表示协调器节点配置更新规则的一般形式。

$\Phi(\cdot)$表示传感器节点的角色配置函数。

$\Phi(h_{j,l,t})$表示在时刻 t 配置联盟 l 的协调器节点 j 的角色。

$B_{i,t+1}$表示一系列与传感器网络特定应用相关的线性组合规则。

θ 表示一个种群中各个体可选择的动作组成的策略空间。

n 表示联盟中传感器节点的个数。

$\boldsymbol{U}$ 表示演化博弈中的效用矩阵。

γ_{ij} 表示演化博弈矩阵中策略 i 对于策略 j 的收益。

$p_i(t)$表示在时刻 t 选择策略 i 的个体比例。

$f_i(t)$表示在时刻 t 策略 i 的自适应度。

$\eta(t)$表示整个种群在时刻 t 的平均自适应度。

$\mathbb{G}$ 表示协调器节点选择博弈。

π 表示参与者的策略空间集合。

π_1 表示防御者 defender 可采取的策略空间。

AC 表示可选协调器节点规则。

TC 表示临时协调器节点规则。

TO 表示临时普通节点。

Φ_{AC}表示使用 AC 规则来配置节点的角色。

Φ_{TC}表示使用 TC 规则来配置节点的角色。

Φ_{TO}表示配置节点的角色为普通节点。

π_2 表示攻击者 attacker 采取的策略空间。

P 表示联盟 M 的传感器节点组合收益。

f'表示参与者的自适应度。

λ 表示选择新的协调器节点的密度参数。

P_j 表示协调器节点选择它的邻居节点 j 并配置其为一个协调器节点的概率。

$h_{i,t+1}^{\mathrm{AC}}$表示使用 AC 规则配置更新协调器节点。

$h_{i,t+1}^{\mathrm{TC}}$表示使用 TC 规则配置更新协调器节点。

$\Phi_{\mathrm{AC}}(h_{j,t})$表示在时刻 t 邻居节点 j 被配置为可选协调器节点的概率。

$\Phi_{\mathrm{TC}}(h_{j,t})$表示在时刻 t 邻居节点 j 被配置为临时协调器节点的概率。

$\Phi_{\mathrm{TO}}(h_{k,t})$表示在时刻 t 原协调器节点 k 被配置为普通节点的概率。

$\overline{M}$ 表示协调器节点的个数。

$c_{\Phi}^{\mathrm{AC}}(i)$表示协调器节点 i 配置为 AC 的成本函数。

$c_{\Phi}^{\mathrm{TC}}(i)$表示协调器节点 i 配置为 TC 的成本函数。

$c_{\Phi}^{\mathrm{TO}}(i)$表示协调器节点 i 配置为 TO 的成本函数。

$c_{\Phi}(\cdot)$表示配置时消耗传感器节点的能量成本。

r_{Φ} 表示具有 $\overline{M}$ 个协调器的传感器节点联盟的配置性能。

w_i 表示权重。

$Q(\cdot)$表示演化状态的配置性能函数。

ω 表示学习率参数。

$\tilde{\omega}$ 表示学习率参数。

ξ 表示打折因子。

Z_t 表示在时刻t 协调器节点选择博弈的配置更新状态。

a_t 表示在时刻t 协调器节点选择和传感器节点配置的行动。

χ 表示协调器节点的可靠状态等级。

C_0 表示协调器节点的信道占用情况。

E_0 表示协调器节点的剩余能量。

$O(\chi)$表示协调器节点的选择等级。

$\mu_A(x)$表示模糊集的成员的隶属度函数。

ϵ 表示模糊成员均值。

σ 表示偏移标准差。

μ_A^* 表示反模糊化输出。

n'表示采样值的个数。

x_i 表示传感器节点的信道和能量的采样值。

$\mu_A(x_i)$表示 x_i 采样的隶属函数值。

$P_i(t)$表示在时刻 t 传感器节点 i 被选择为协调器节点的概率。

$p_{i,a_i}(t)$表示在时刻 t 传感器节点被选择为协调器节点并复制策略和配置更新规则的概率。

7.2　相关工作

ZigBee 无线传感器网络安全和保证 QoS 相结合的无线传感器网络部署和管理是目前较为活跃的研究领域。由于 ZigBee 无线传感器网络数据传输受能量约束，因此如何将数据

包以能量高效的方式路由到目标节点，并延长网络生命期是当前面临的一个挑战问题[362]。大量的文献已研究了ZigBee无线传感器网络中的路由协议[363-365]，其中，分层协议[366, 367]把节点分成了传感簇[368-371]。在文献[372]中还研究了单个数据包的路由跳数和网络能量。然而，这些协议未考虑协调器节点受到恶意节点攻击时的防御策略。

ZigBee无线传感器网络的安全问题一直备受关注。通常在无线通信网络中，应用高效的基于散列链的轻量级认证协议可以防御针对中继节点的攻击，典型的协议有计时流损失容错认证协议，其实质是一个基于宽松时间同步的广播认证协议[373]。Law等人[374]阐述了如何使用干扰实现对网络链接层的攻击。Xu等人[375]通过对攻击和防御的测试研究实现了对传感器网络的攻击，他们提出的干扰攻击模型包括恒定干扰、欺骗性干扰、随机干扰和反应式干扰。Yao等人[376]为无线传感器网络安全提出了一个信任管理机制，它实际上是一个安全路由协议。在这个安全路由协议中，每个节点使用参数评估邻居节点。Aivaloglou和Gritzalis[132]提出了基于证书和行为评估的混合信任/信誉管理协议。Gabrielli等人[377]分析了典型拓扑协议PEAS、ASCENT和CCP的安全漏洞，并重新设计了相应的安全拓扑协议。Bao等人[130]结合社会网络属性等多维的信任属性来评估每个传感器节点的信任度。Zonouz等人[100]使用博弈论把对攻击者的防御模型化为两个参与者的Stackelberg随机博弈，并且使用模糊逻辑推理计算网络层安全测量值。

不同的ZigBee无线传感器网络拓扑对其安全性有重要影响。除了干扰攻击外，ZigBee无线传感器网络主要还有三类攻击：第一类是针对ZigBee组件和配置的漏洞攻击；第二类是窃听ZigBee网络中加密或未加密的数据，从而获得与用户相关的敏感性信息；第三类是重传捕获的数据。对于第一类攻击，可以通过减少网络配置，提高设备发现协议的认证效率来降低漏洞的泄露机会。对于第二类攻击，可通过提高数据传输的保密性来应对。第三类攻击主要是使用重传方式消耗节点的能量，导致数据传输链接失败，因此，针对此类攻击的一般防御方案是，不论受到攻击与否，配置ZigBee节点为Sleep-wake周期性工作方式，使得其在完成传输任务前，能量耗尽情况尽量不会发生[378]。Patel等人[379]针对传统的MDA-ML方法导致ZigBee设备认证性能降低的问题，提出非参数随机森林(Non-parametric Random Forest)和多层次演算分类器(Multi-Class AdaBoost)算法来提高ZigBee设备的认证性能。通常，ZigBee安全协议在发送数据前先使用高级加密的标准计数器模式加密数据，这种模式中过多的异或操作消耗了时间，减少了数据传输的实时性。针对该问题，Bakhache等人[380]为了提高加密算法的健壮性和实时性，利用混沌函数高效的加密性能，提出了快速混沌加密算法(Robust and Fast Chaotic Encryption Algorithm)。典型的ZigBee无线传感器网络密钥管理机制有椭圆曲线Diffie-Hellman(ECDH)机制，但它并不能抵御中间人攻击，在文献[381]中，Choi等人针对中间人攻击问题，结合ECDH和SubMAC(Sub Message Authentication Code)管理机制，增强了ZigBee无线传感器网络中密钥的管理。由于ZigBee无线传感器网络具有低功耗约束，在众多的加密算法中，并不是所有的算法都适用于ZigBee无线传感器网络的数据加密，Rosli等人[382]在比较分析了各种加密算法的性能后，得出IBE-Trust协议比RSA-1024消耗更少能量的结论。Xu等人[383]使用动态密钥及密钥的同步更新技术，设计了wz-lcp协议来满足智能家庭应用中的低功耗和高安全性需求。具有128位加密密钥的ZigBee无线传感器网络，也容易被攻击者截获，需要跨层设计其安全性，Ramsey等人[384]利用RF物理层特性，设计了PHY-MAC-NWK跨层安全框架。

Jokar 等人[385]通过在家庭网络中部署欺骗探测模块和在网络节点上部署阻止模块，根据接收到帧中的信号强度过滤恶意帧。Tseng 等人[386]使用随机博弈对异构网络的网络认证问题建立了相应的模型。Jiang 等人[387]使用演化博弈模型分析信息的扩散过程和自适应网络的滤波问题。

与上述相关工作相比，本章使用演化博弈和随机博弈建立 ZigBee 无线传感器网络安全防御模型，当协调器节点受到攻击时，给出相应的响应策略。与文献[386]相比，本章使用博弈论解决了 ZigBee 无线传感器网络协调器节点受攻击响应问题。与文献[387]相比，本章主要针对协调器节点受到攻击时，使用演化博弈模型分析协调器节点选择问题，并且结合模糊理论，运用随机演化博弈和模糊推理给出在安全和 QoS 约束下的基于协调器节点选择的协作防御恶意行为的方法。

7.3　系统模型

7.3.1　ZigBee 无线传感器网络的功能性和 QoS

传感器节点提供的功能包括数据包接收和转发操作。当数据包到达传感器节点后，传感器节点接收该数据包再转发给协调器节点，然后网络中的传感器节点聚合后提供组合服务给终端节点。与具有丰富带宽和较高可用性的 Web 服务提供者不同，传感器网络有很高的动态性。由于传感器节点容易失效，通信链接容易断开，使得无线通信容量受到限制。因此，ZigBee 无线传感器网络的稳定运行包含两个 QoS 属性，即从终端节点到协调器节点的平均跳数和能量成本。其中，跳数定义为覆盖了源节点到协调器节点总的路径长度，能量成本定义为 ZigBee 无线传感器网络消费者偿付接收和转发数据包操作的费用。

这里利用演化联盟博弈分析协调器节点的选择过程，其拓扑结构如图 7-1 所示。在图 7-1 中，簇联盟能够通过协调器节点与其他联盟通信。每个联盟包含协调器节点和多个传感器节点等联盟成员。协调器节点负责接收和转发数据包并且管理联盟成员间的合作，它不但可以请求或被邀请加入联盟，也可以退出联盟。

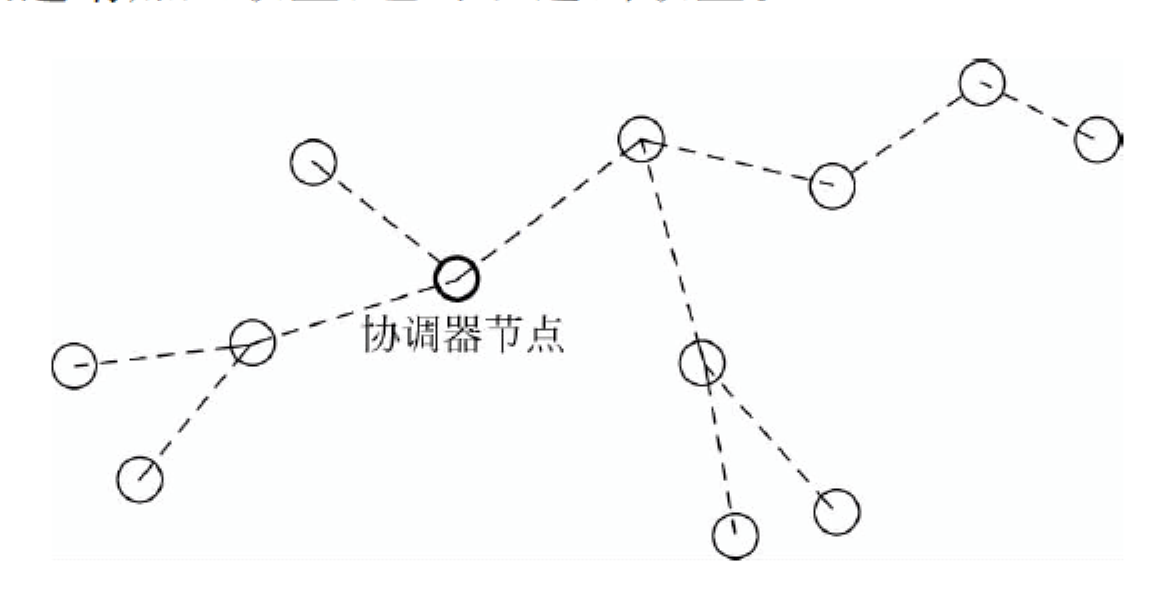


图 7-1　协调器节点选择的联盟博弈拓扑

在一个联盟中，协调器节点是联盟的控制器，它负责初始化网络设置，选择合适的通信信道。通过对联盟内可用信道的能量峰值来评估信道间的干扰。如果一个节点是可用的，将开始和协调器节点交互数据包形成联盟，同时，协调器节点将选择其作为自己的联盟成员。合作的联盟是由协调器节点管理的传感器节点组合，记为 $S=[s[1],\cdots,s[n]]$，其中具有 n 个联盟成员。一旦协调器节点接收到一个传感器节点的数据发送请求后，它将转发给

联盟成员节点。如果 $s[1]$ 节点是可用的，协调器节点将发送数据给 $s[1]$；否则将发送给 $s[2]$，……，直到数据被转发到目标节点。假设联盟传感器节点 $s[i]$ 的可用性是 av_i，则整个联盟 M 的可用性为

$$\mathrm{av}_M = 1 - \prod_{i=1}^{n}(1-\mathrm{av}_i) \tag{7-1}$$

由于联盟是一个组合传感器网络，使用每个传感器节点时将产生的费用记为 p_M。组合传感器网络的成本决定于平均跳数、能量成本，因此，联盟 M 的成本为

$$c_M(s) = \sum_{i=1}^{n}\prod_{j=1}^{i}(1-\mathrm{av}_{s[j]})c_{s[i]} \tag{7-2}$$

式中，$c_{s[i]}$ 为节点 $s[i]$ 加入联盟的成本。令 h_i 为从联盟节点 $s[i]$ 到协调器节点的平均跳数，则单个数据包花费的能量成本为

$$c_{s[i]} = (E_r + E_f) \times h_i \tag{7-3}$$

式中，E_r 和 E_f 分别为接收和发送单个数据包的能量成本。所以，联盟 M 的传感器节点组合成本为

$$v_M(s) = p_M \times \mathrm{av}_M - c_M(s) \tag{7-4}$$

7.3.2 协调器节点攻击响应的随机博弈模型

这里将协调器节点攻击问题看作是一个非合作博弈问题。其中，恶意节点为攻击者，协调器节点为防御者，恶意节点攻击的结果是使得协调器节点工作状态发生变化。协调器节点有两个内部运行状态，即 NormalState(NS)、HackedState(HS)。攻击者通过改变其内部运行状态使其失去对联盟的控制能力，博弈模型表示如下：

定义 7-1 协调器节点攻击响应的随机博弈模型是一个四元组，即

$$\widetilde{G} = (N, Z, \{A_k \mid k \in N\}, \{u_k \mid k \in N\}) \tag{7-5}$$

式中，N 为博弈参与者集合，$N=\{1,2\}=\{\text{attacker}, \text{defender}\}$；$Z$ 为状态空间，$Z=\{\text{NS}, \text{HS}\}$；$\{A_k|k\in N\}$ 为参与者 k 采取的行动集合；$A_1=\{a_1, r_1\}$，其中 a_1 表示使协调器节点状态从 NormalState 改变为 HackedState 状态的攻击行动；r_1 表示攻击者在 NormalState 状态实施的攻击行动；$A_2=\{d_2, \varnothing_2\}$，其中，$d_2$ 表示攻击行动 a_1 被防御者成功检测，$\varnothing_2$ 表示攻击行动未被防御者 defender 成功检测；$\{u_k|k\in N\}$ 为参与者 k 的效用函数。

防御者 defender 的目标是通过协调器节点选择调度机制最大化延长网络的生命期，而恶意的攻击者 attacker 通过干扰策略降低网络生命期。因此，防御者 defender 和攻击者 attacker 有相反的目标，其交互过程可以动态模型化为一个非合作零和博弈。在协调器节点选择调度机制中，协调器节点作为一个理性的博弈防御者，它的目标是最大化网络吞吐量、降低数据包传输的平均跳数和能量成本。因此，定义防御者 defender 的效用函数为防御行动所获得的期望收益，即

$$u_2(a_2, a_{-2}) = \sum_{a\in A_2} p(a)u_2(a) \tag{7-6}$$

式中，a_2 为防御者 defender 采用的防御行动；a_{-2} 为攻击者 attacker 采取的攻击行动；$p(a)$ 为防御者 defender 在其行动空间 A_2 中采取防御行动的概率；$u_2(a)$ 为防御者 defender 在其行动空间 A_2 中采取防御行动后，成功检测攻击获得的收益。防御者 defender 在其行动空

间 A_2 中，为最大化防御收益而选择最优响应行动。因此，在博弈中防御者 defender 的决策为

$$(\widetilde{G})：\max\ u_2(a_2, a_{-2}) \tag{7-7}$$

7.3.3　基于演化博弈的最优响应策略

定义协调器节点选择的联盟博弈拓扑为一个有向图 $G=(\widetilde{N}, E)$，其中 $\widetilde{N}$ 是节点集合，E 是边集。假设每个节点初始的资源可用。令 r_i 是联盟节点 $s[i]$ 防御攻击的容侵性，x_i 和 y_i 分别是分配给防御者 defender 和攻击者 attacker 的资源。根据竞争模型[388]思想，联盟节点 $s[i]$ 的资源可用性的损失率为

$$\tau_i(x_i) = \frac{(y_i)^m}{\alpha_i\ (r_i + x_i)^m + (y_i)^m} \tag{7-8}$$

式中，$m\in(0,1]$ 为联盟节点 $s[i]$ 资源可用性的线性和非线性损失率；α_i 为防御者 defender 保护协调器节点的困难程度。当 $\alpha_i\in(0,1)$ 时，防御者 defender 比攻击者 attacker 分配更多的资源来缓解攻击效应。当 $\alpha_i>1$ 时，意味着防御者 defender 可以成功地检测或缓解攻击。这样，可重写联盟 M 的传感器节点组合成本为

$$v'_M(s) = p_M \times \mathrm{av}_M - c_M(s) - \tau_M(s) \tag{7-9}$$

式中，$\tau_M(s)$ 为联盟 M 的可用资源损失率，定义为 $\sum_{i=1}^{n}\tau_i(x_i)$。对于防御者 defender 而言，它的目标是使用资源 x_i 来最大化 $v'_M(s)$，并且通过选择协调器节点最大化网络的 QoS。攻击者 attacker 的目标是使用资源 y_i 攻击关键的协调器节点来最小化 $v'_M(s)$。从式(7-9)中可以看出，如果 $c_M(s)$ 减少，则 $v'_M(s)$ 增大。同时还注意到，$c_M(s)$ 的值依赖于 h_i、E_r 和 E_t，当这 3 个变量中至少有一个变量减少时，$c_M(s)$ 才减少。其中 h_i 能够通过演化选择协调器节点位置和可靠状态获得。对于具有度为 d_i 的协调器节点，联盟集合可以表示为 $\{i_1,\cdots,i_{d_i}\}$，这样协调器节点配置更新规则的一般形式可写为

$$h_{i,t+1} = \sum_{l=1}^{d_i}\sum_{j=1}^{n} B_{i,t+1}(j,l)\Phi(h_{j,l,t}) \tag{7-10}$$

式中，$\Phi(\cdot)$ 为传感器节点的角色配置函数；$\Phi(h_{j,l,t})$ 为在时刻 t 配置联盟 l 的协调器节点 j 的角色；$B_{i,t+1}$ 为一系列与传感器网络特定应用相关的线性组合规则。

演化博弈论(EGT)来源于生态生物学的研究[389]，它主要强调个体策略的动态性和稳定性。记 $\theta=\{\theta_1,\theta_2,\cdots,\theta_k\}$ 为一个种群中各个体可选择的动作组成的策略空间。效用矩阵为 $\boldsymbol{U}$，用一个 $k\times k$ 的矩阵表示，其中矩阵的元素 γ_{ij} 表示策略 i 对于策略 j 的收益。在时刻 t 选择策略 i 的个体比例用 $p_i(t)$ 来表示，且 $0<p_i(t)<1$，$i\in\{1,\cdots,k\}$。在时刻 t 策略 i 的自适应度为

$$f_i(t) = \sum_{j=1}^{k} p_j(t)\gamma_{ij} \tag{7-11}$$

整个种群在时刻 t 的平均自适应度为

$$\eta(t) = \sum_{i=1}^{k} p_i(t) f_i(t) \tag{7-12}$$

每个参与者的策略更新方程为

$$p_i(t+1)=\frac{p_i(t)f_i(t)}{\eta(t)} \tag{7-13}$$

实际上，演化博弈中的策略更新过程类似于协调器节点位置选择及配置更新过程。因此，定义协调器节点选择博弈如下。

定义 7-2 协调器节点选择博弈是一个三元组$G=(N',\pi,P)$，其中：

- $N'=\{\text{attacker},\text{defender}\}$表示参与者集合。
- $\pi=\{\pi_1,\pi_2\}$表示参与者的策略空间集合，$\pi_1=\{AC,TC,\Phi_{AC},\Phi_{TC},\Phi_{TO}\}$表示防御者 defender 可采取的策略空间，其中，AC 表示可选协调器节点规则，TC 表示临时协调器节点规则，Φ_{AC}表示使用 AC 规则来配置节点的角色，Φ_{TC}表示使用 TC 规则来配置节点的角色，Φ_{TO}表示配置节点的角色为普通节点，π_2 表示攻击者 attacker 采取的策略空间。
- $P=v'_M(s)$表示联盟 M 的传感器节点组合成本。

在 EGT 中，每个协调器节点代表一个防御参与者，参与者的自适应度 f'通过与局部的邻接参与者交互来决定，它以分布式自适应选择协调器节点的方式来更新协调器节点的配置，其表达式为

$$f'=(1-\lambda)v_M(s)+\lambda v'_M(s) \tag{7-14}$$

式中，λ 为选择新的协调器节点的密度参数。当 $\lambda\to 0$ 时表示弱的干扰攻击，重新选择协调器节点的概率较小；$\lambda=1$ 表示强选择，重新选择协调器节点的概率较大。接下来，使用规则来描述协调器节点选择的动态性。

规则 7-1 可选协调器节点更新规则(AC 规则)。联盟中的协调器节点放弃当前作为协调器节点的角色，随后，再以一定概率选择它的邻居节点作为协调器节点，之后，邻居节点复制其策略并配置为协调器节点，具体过程如图 7-2 所示。

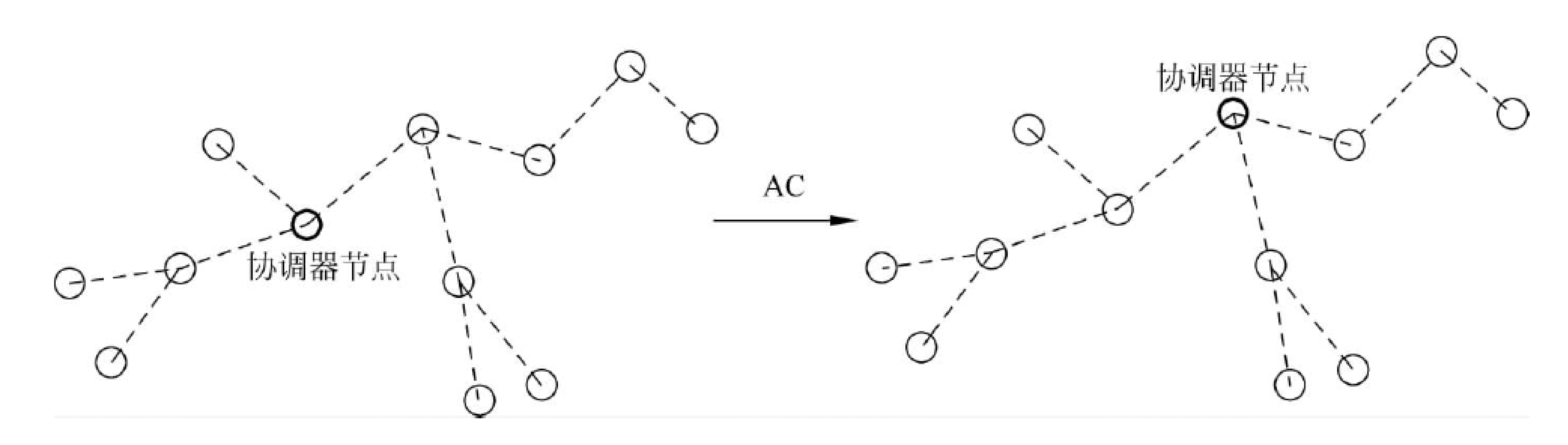


图 7-2 可选协调器节点更新规则

规则 7-2 临时协调器节点更新规则(TC 规则)。当邻居节点采用该规则时，将被配置为临时协调器节点，而原协调器节点被配置为一个临时的普通节点 TO(Temporary Ordinary)。当干扰攻击变弱时，邻居节点和原协调器节点恢复其原来的角色，这个过程如图 7-3 所示。

这两种更新规则能以自适应的方式匹配协调器节点配置更新算法。对于可选协调器节点更新规则，协调器节点选择它的邻居节点 j 并配置其为一个协调器节点的概率为

$$P_j=\frac{f_j}{\sum_{q=1}^{n}f_q}\cdot\frac{1}{\Gamma_j} \tag{7-15}$$

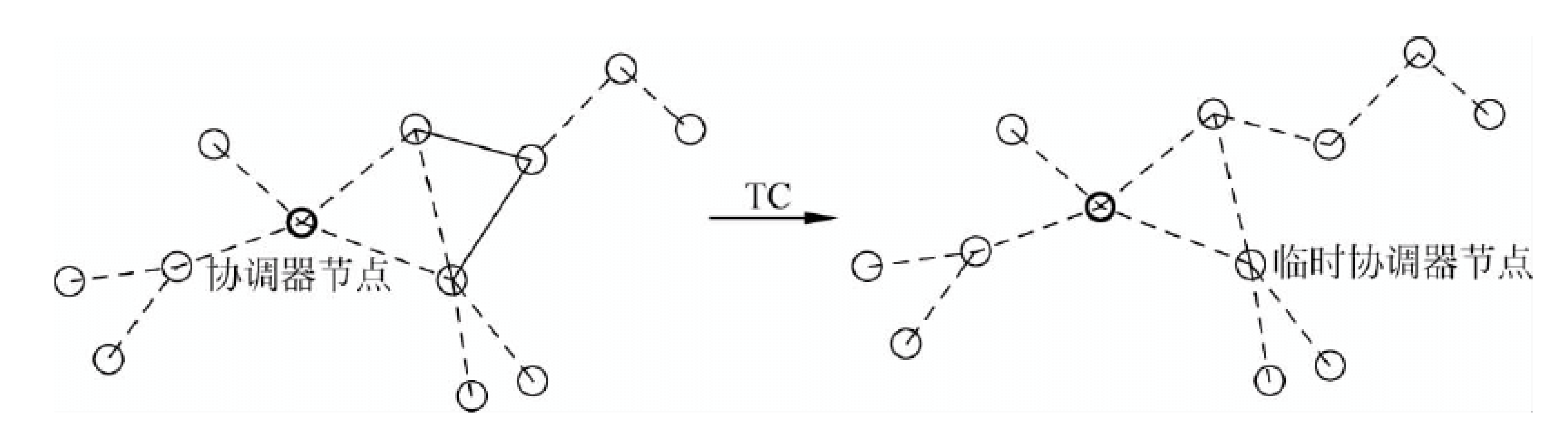


图 7-3　临时协调器节点更新规则

式中，$\frac{f_j}{\sum_{q=1}^{n} f_q}$ 为邻居节点 j 被选择成为协调器节点的概率；f_j 为邻居节点 j 的自适应度；$1/\Gamma_j$ 为邻居节点 j 复制协调器节点策略及其配置的概率。对于可选协调器节点更新规则，协调器节点配置更新规则形式化表示为

$$h_{i,t+1}^{\mathrm{AC}} = \left(\frac{f_j}{\sum_{q=1}^{n} f_q} \cdot \frac{1}{\Gamma_j}\right)\Phi_{\mathrm{AC}}(h_{j,t}) + \left(1 - \frac{f_j}{\sum_{q=1}^{n} f_q} \cdot \frac{1}{\Gamma_j}\right)\sum_{i \in M\backslash\{j\}} \Phi(h_{i,t}) \tag{7-16}$$

式(7-16)第一部分表示邻居节点配置为一个 AC 规则的概率，其中，$\Phi_{\mathrm{AC}}(h_{j,t})$表示在时刻 t 邻居节点 j 被配置为可选协调器节点的概率；第二部分是联盟中余下所有节点更新其到目标节点路径的平均跳数的概率，其中，$i \in M\backslash\{j\}$表示联盟中余下的所有节点。

对于临时协调器节点更新规则，协调器节点重新选择后的配置更新规则形式化表示为

$$h_{i,t+1}^{\mathrm{TC}} = \left(\frac{f_j}{\sum_{q=1}^{n} f_q} \cdot \frac{1}{\Gamma_j}\right)\Phi_{\mathrm{TC}}(h_{j,t}) + \left(\frac{f_j}{\sum_{q=1}^{n} f_q} \cdot \frac{1}{\Gamma_j}\right)\Phi_{\mathrm{TO}}(h_{k,t}) + \left(1 - \frac{f_j}{\sum_{q=1}^{n} f_q} \cdot \frac{1}{\Gamma_j}\right)\sum_{i \in M\backslash\{j,k\}} \Phi(h_{i,t}) \tag{7-17}$$

式(7-17)第一部分表示邻居节点配置为一个临时协调器节点的概率，其中，$\Phi_{\mathrm{TC}}(h_{j,t})$表示在时刻 t 邻居节点 j 被配置为临时协调器节点的概率；第二部分表示原协调器节点本身配置为一个临时的普通节点的概率，其中，$\Phi_{\mathrm{TO}}(h_{k,t})$表示在时刻 t 原协调器节点 k 被配置为普通节点的概率；第三部分表示联盟中余下所有节点更新其到目标节点路径的平均跳数的概率，其中，$i \in M\backslash\{j,k\}$表示联盟中余下的所有节点。

由于可选协调器节点更新规则和临时协调器节点更新规则在配置网络时需消耗传感器节点的能量成本，因此，联盟 M 的传感器节点组合的收益重写为

$$v''_M(s) = p_M a_M - c_M(s) - \tau_M(s) - c_\Phi(s) \tag{7-18}$$

式中，$c_\Phi(\cdot)$为配置时消耗传感器节点的能量成本。对于具有 $\bar{M}$ 个协调器的传感器节点联盟的配置性能为

$$r_\Phi = \sum_{i=1}^{\bar{M}} (w_i c_\Phi^{\mathrm{AC}}(i) + (1 - w_i)(c_\Phi^{\mathrm{TC}}(i) + c_\Phi^{\mathrm{TO}}(i))) \tag{7-19}$$

式中，w_i 为权重，且 $w_i < 1$。$c_\Phi^{\mathrm{AC}}(i)$为协调器节点 i 配置为 AC 的成本函数；$c_\Phi^{\mathrm{TC}}(i)$为协调器节点 i 配置为 TC 的成本函数；$c_\Phi^{\mathrm{TO}}(i)$为协调器节点 i 配置为 TO 的成本函数。演化状态的

配置性能由函数 $Q(\cdot)$测量，这里使用即时差分(Temporal Difference，TD)法更新 $Q(\cdot)$函数值。TD 方法的基本思想是只要观察到有收益，不必等待输出最终的收益，就能够与前一状态的收益进行差分，从而缩短了更新时间。$Q(\cdot)$函数更新表达式为

$$Q(Z_t,a_t) \leftarrow Q(Z_t,a_t) + \omega[r_{\Phi,t+1} + \xi Q(Z_{t+1},a_{t+1}) - Q(Z_t,a_t)] \tag{7-20}$$

式中，ω 为学习率，在协调器节点受到攻击时，保证协调器节点选择博弈能够收敛；ξ 为打折因子，在连续的配置更新任务中，保证累积的收益能够收敛；Z_t 为在时刻 t 协调器节点选择博弈的配置更新状态；a_t 为在时刻 t 协调器节点选择和传感器节点配置的行动。

7.4 基于 FQL 增强学习的协调器节点选择

7.4.1 模糊逻辑

本节使用 Fuzzy-Q 学习(FQL)算法来实现演化博弈决策(FEGD)过程。FEGD 方法考虑协调器节点的信道占用情况 C_0 和剩余的能量 E_0，它们之间的相关度用函数 $\chi = f(C_0, E_0)$表示，其中，χ 表示协调器节点的可靠状态等级。在 FEGD 方法中，输入参数为协调器节点已占用的信道 C_0 和剩余的能量 E_0。输入参数的集合分别定义为

$$T(C_0) = \{\text{Low(LO)}, \text{High(HG)}\} \tag{7-21}$$

$$T(E_0) = \{\text{Low(LO)}, \text{Moderate(ME)}, \text{High(HG)}\} \tag{7-22}$$

输出参数为协调器节点的选择等级，定义为

$$O(\chi) = \{\text{Low(LO)}, \text{Moderate(ME)}, \text{High(HG)}\} \tag{7-23}$$

表 7-1 给出了模糊规则矩阵。按照 FQL 算法，为演化博弈定义了包含有 4 条规则的模糊推理系统：

(1) IF C_0 is HG AND E_0 is HG THEN χ is HG.

(2) IF C_0 is HG AND E_0 is ME THEN χ is ME.

(3) IF C_0 is LO AND E_0 is ME THEN χ is ME.

(4) IF C_0 is LO AND E_0 is HG THEN χ is ME.

表 7-1 模糊规则矩阵

E_0 \ C_0	LO	HG
LO	LO	LO
ME	ME	ME
HG	ME	HG

选用高斯模糊成员函数，定义模糊集的成员的隶属度函数为

$$\mu_A(x) = \exp\left(-\frac{(x-\varepsilon)^2}{2\sigma^2}\right) \tag{7-24}$$

式中，x 为连续的信道占用情况和剩余的能量采样值；ε 为模糊成员均值；σ 为偏移标准差。$\mu_A(x): U \to [0,1]$表示连续取值的特征函数，即隶属函数。使用重心法反模糊化，反模化的结果为

$$\mu_A^* = \frac{\sum_{i=1}^{n'} \mu_A(x_i) \times x_i}{\sum_{i=1}^{n'} \mu_A(x_i)} \tag{7-25}$$

式中，n'为采样值的个数；x_i 为传感器节点的信道和能量采样值；$\mu_A(x_i)$为 x_i 采样的隶属函数值。

7.4.2　随机学习过程

使用FQL算法能够推导出传感器节点的信道和能量情况，为协调器节点的选择博弈提供决策依据，但是在博弈过程中，攻击者的策略是随时间变化的，多个参与者可同时采取防御行动，通过分布式的随机演化学习算法(SEL)，使得协调器节点的选择博弈达到纳什均衡。为了方便开发基于SEL的协调器节点选择的自组织防御算法，令混合策略 $P_i(t)=[p_{i,1}(t),\cdots,p_{i,|\pi_1|}(t)]$表示在时刻 t 传感器节点 i 被选择为协调器节点的概率。其中，$p_{i,a_i}(t),a_i\in\{1,2,\cdots,|\pi_1|\}$，表示在时刻 t 传感器节点被选择为协调器节点并复制策略和配置更新规则 $a_i\in\pi_1$ 的概率。基于SEL的协调器节点选择形成自组织防御的算法如算法7-1所示。

算法7-1　基于SEL的协调器节点选择形成自组织防御的算法(SoDSC)。

1. 初始化设置 $t=0$，传感器节点被选择为协调器节点的概率 $p_{i,a_i}(t)=1/\Gamma_j$。
2. 在时刻 t，每个参与者 i 选择一个行动 $a_i(t)$进行防御。
3. 联盟接收收益 $v_M(s)$。
   //在联盟中的每个传感器节点根据以下规则更新被选择为协调器节点的概率
4. CID=Getcurrent_node(ID)。
5. CRS=Get_CoordinatorResourceState(CID)。
6. REPEAT
7. 　IF CRS= HG
8. 　　$h_{i,t+1}\leftarrow h_{i,t}$；
9. 　　$p_{i,a_i}(t+1)\leftarrow p_{i,a_i}(t)+\tilde{\omega}h_{i,t+1}(\mathbf{1}_{\{a_i=\mathrm{AC}\vee\mathrm{TC}\}}-p_{i,a_i}(t))$。
   　　//$0<\tilde{\omega}<1$ 表示学习率，$\mathbf{1}_{\{.\}}$表示指示函数
   　　//$a_i=\mathrm{AC}\vee\mathrm{TC}$ 表示防御者采取AC或TC规则配置时，指示函数值为1；否则//为0
10. 　　$Q(Z_t,a_t)\leftarrow Q(Z_t,a_t)+\omega\left[\sum_{i=1}^{M}((w_i\cdot c_{\Phi}^{\mathrm{AC}}(i)+\xi\cdot Q(Z_{t+1},a_{t+1})\cdot 1_{\{a_i=\Phi_{\mathrm{AC}}\}}-Q(Z_t,a_t))\right]$。
   　　//$a_i=\Phi_{\mathrm{AC}}$表示防御者采用AC规则配置时，指示函数 $1_{\{.\}}$的值为1；否则为0
11. 　ELSEIF CRS= ME
12. 　　$h_{i,t+1}\leftarrow h_{i,t}$。
13. 　　$p_{i,a_i}(t+1)\leftarrow p_{i,a_i}(t)+\tilde{\omega}h_{i,t+1}(\mathbf{1}_{\{a_i=\mathrm{AC}\vee\mathrm{TC}\}}-p_{i,a_i}(t))$。
14. 　　$Q(Z_t,a_t)\leftarrow Q(Z_t,a_t)$

$$+\omega\left(\left(\sum_{i=1}^{\bar{M}}((1-w_i)(c_{\Phi}^{\mathrm{TC}}(i)+c_{\Phi}^{\mathrm{TO}}(i)))+\xi\cdot Q(Z_{t+1},a_{t+1})\cdot \mathbf{1}_{\{a_i=\Phi_{\mathrm{TC}}\wedge\Phi_{\mathrm{TO}}\}}-\right.\right.$$

$$\left.Q(Z_t,a_t)\right)。$$

//$a_i=\Phi_{\mathrm{TC}}\wedge\Phi_{\mathrm{TO}}$表示邻居节点被配置为临时协调器节点,并且原协调器节点
//配置为临时的普通节点时,指示函数 $\mathbf{1}_{\{\cdot\}}$ 的值为 1;否则为 0

15. ENDIF

16. UNTIL 函数 $p_{i,a_i}(t+1)$和 $Q(Z_t,a_t)$的值收敛。

算法 7-2 Get_CoordinatorResourceState(CID)

1. 初始化 C_0、E_0。
2. 给定 $T(C_0)$,$T(E_0)$,使用式(7-24)计算 $\mu_A(C_0)$和 $\mu_A(E_0)$。
3. 使用表 7-1 计算 $O(\chi)$。
4. 计算 $\mu_A(O(\chi))=\min(\mu_A(C_0),\mu_A(E_0))$。
5. 使用式(7-25)计算反模糊输出 $\chi=f(C_0,E_0)=\mu_A^*$。
6. 返回 $O(\chi)$。

在算法 7-2 中,协调器节点可以读取电池能量级别并且使用 ACK 消息估计信道状况,使用这些参数,从单个协调器节点选择的角度描述了基于模糊逻辑的协调器节点资源状态决策算法。

7.5 实验

使用网络仿真器 NS-2,仿真实现 IEEE 802.15.4 物理层和 MAC 层标准。首先仿真了动态防御和响应策略,仿真结果显示,通过选择未被攻击的协调器节点形成的新联盟增长了网络的吞吐量。然后,验证了当协调器节点受到攻击且失效时,使用模糊逻辑和演化博弈的响应策略来配置 IEEE 802.15.4/ZigBee 网络,仿真结果显示,提出的算法延长了网络生命期。仿真的网络场景由 20 个传感器节点组成,随机地部署在方形区域。仿真参数如表 7-2 所示。

表 7-2 仿 真 场 景

参　　数	值
Protocols	AODV,Mac/802.15.4
Number of Nodes	20
Simulation Area	50×50
Traffic Type	cbr, poisson
Packet Size	70Bytes
Packets Rate	250k
Distance	5m、9m、10m、11m、12m
Simulation time	100s

NS-2 仿真器实现了 IEEE 802.15.4 标准描述的 Two-way 传播模型,物理层中每个数据包的接收功率应满足一定的阈值,在实验中设定为 3.24×10^{-10}W。仿真开始时随机地选

择协调器节点，通过变换协调器节点的位置和干扰攻击点，对 IEEE 802.15.4 标准和本章提出的算法进行了重复实验。仿真结束时，记录了协调器节点受攻击时新网络拓扑的吞吐量。图 7-4 给出了协调器节点受攻击时，IEEE 802.15.4 标准与协调器节点选择算法获得的吞吐量变化情况。图 7-5 给出了协调器节点受攻击时防御的时间延迟情况。

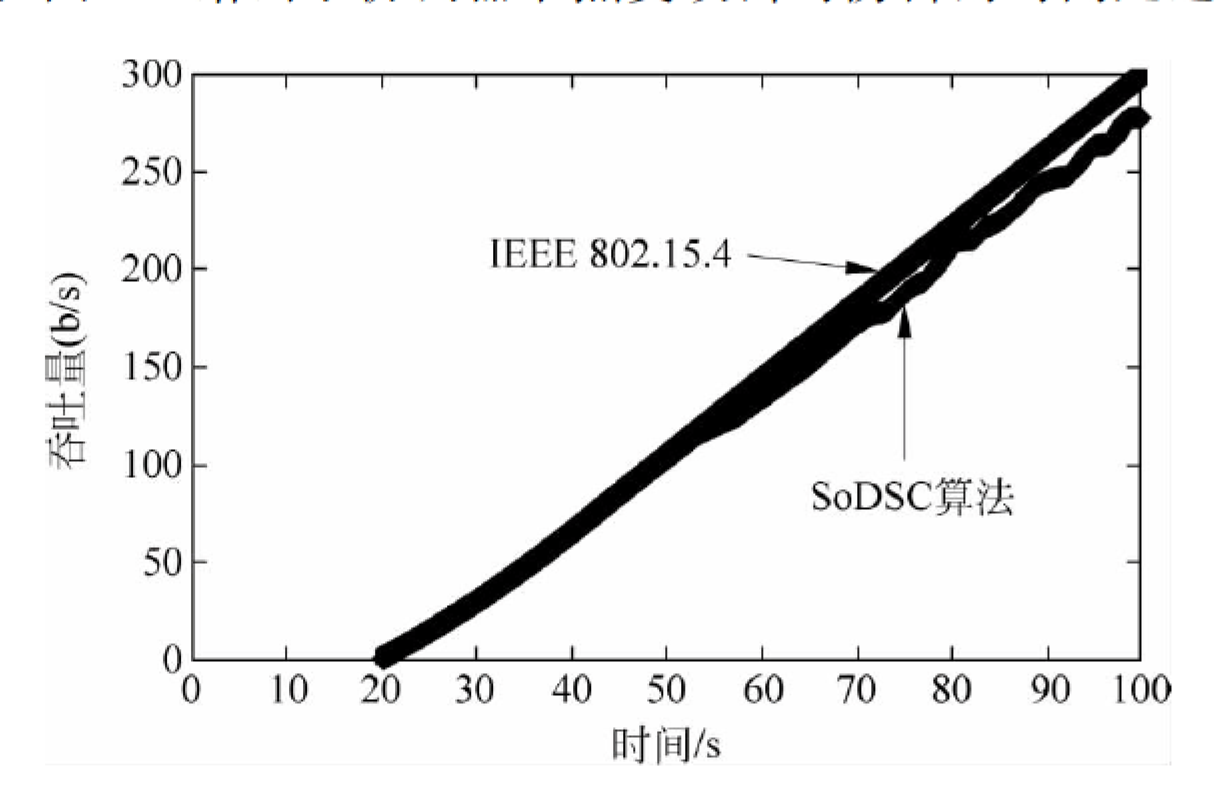


图 7-4　协调器节点被攻击时的网络吞吐量

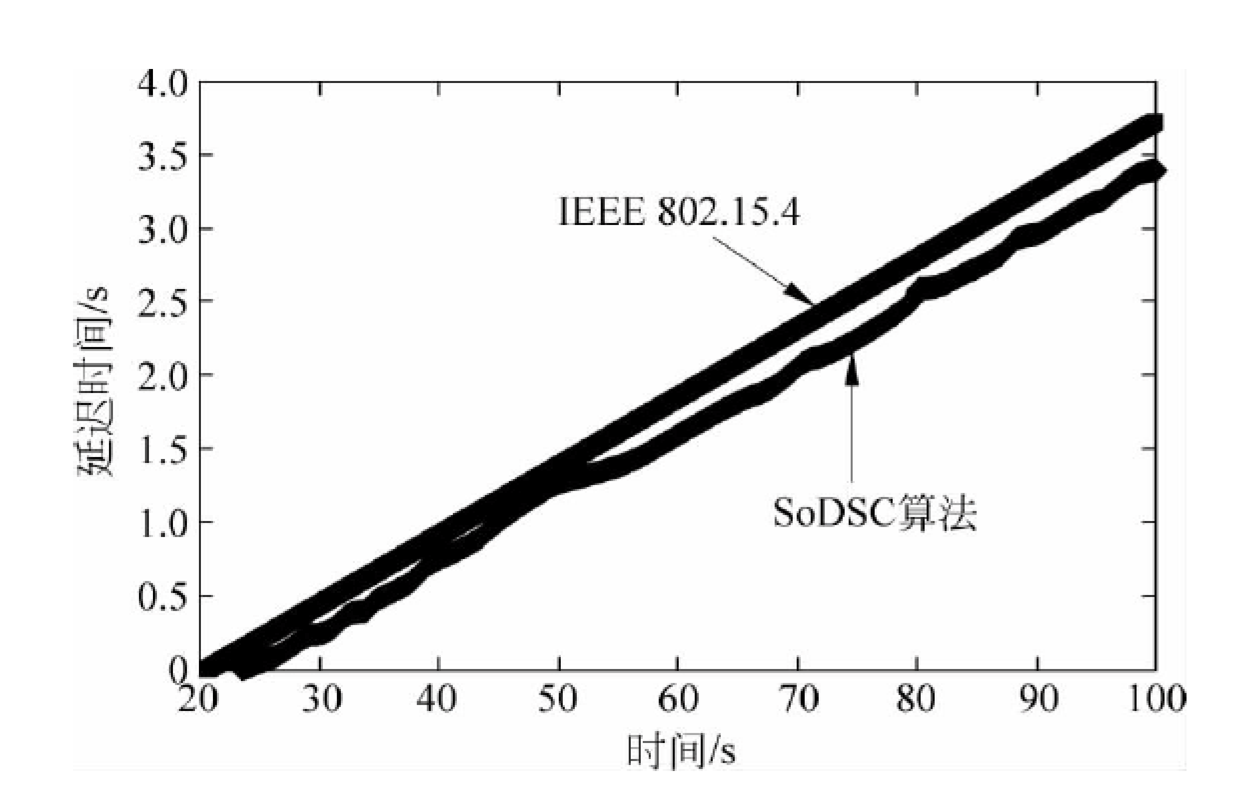


图 7-5　协调器节点被攻击时的防御延迟情况

当网络受到高强度的干扰攻击时，通过博弈选择协调器节点，并控制拓扑来完成干扰攻击的防御，保证了受攻击网络的吞吐量和延迟与 IEEE 802.15.4 相当。这是由于在面对攻击时，博弈选择算法使用 TC 和 AC 规则选择协调器节点来保持较高的吞吐量。同时，在源节点和协调器节点之间创建了较短的路径，使得数据传输的平均跳数减小，节约了能量和减少了数据分发延迟。

在面对干扰攻击时协调器节点选择的推理系统如图 7-6 和图 7-7 所示。图 7-6 给出了协调器节点选择的成员函数曲线，其中包含节点信道质量的成员函数曲线、节点能量状况的成员函数曲线、选择等级的成员函数曲线。通过观察节点信道质量的成员函数曲线可以看出，当节点信道模糊值约为 10 时，信道质量隶属值最高，说明此时节点的信道质量最好。当节点信道模糊值约为 5 时，信道质量隶属值为中等，说明此时节点具有中等的信道质量。通过观察节点能量状况的成员函数曲线可以发现，当节点能量模糊值约为 0.1 时，节点能量较低，当节点能量模糊值约为 10 时，节点能量最高。从选择等级的成员函数曲线可以看出，当选择等级模糊值在[9.5 10.5]区间时，节点被选择为协调器节点的概率较高；当选择等级

模糊值在[2 4]区间时，节点被选择为协调器节点的概率为中；当选择等级模糊值在[0 0.3]区间时，节点被选择为协调器节点的概率较低。图 7-7 给出了协调器节点选择模糊推理曲面。当节点信道模糊值约为 10、节点能量模糊值约为 5 时，从模糊推理曲面图中可以得到节点的选择等级模糊值约为 2.1，由图 7-6 的观察结果可以得出节点被选择为协调器节点的概率为中。

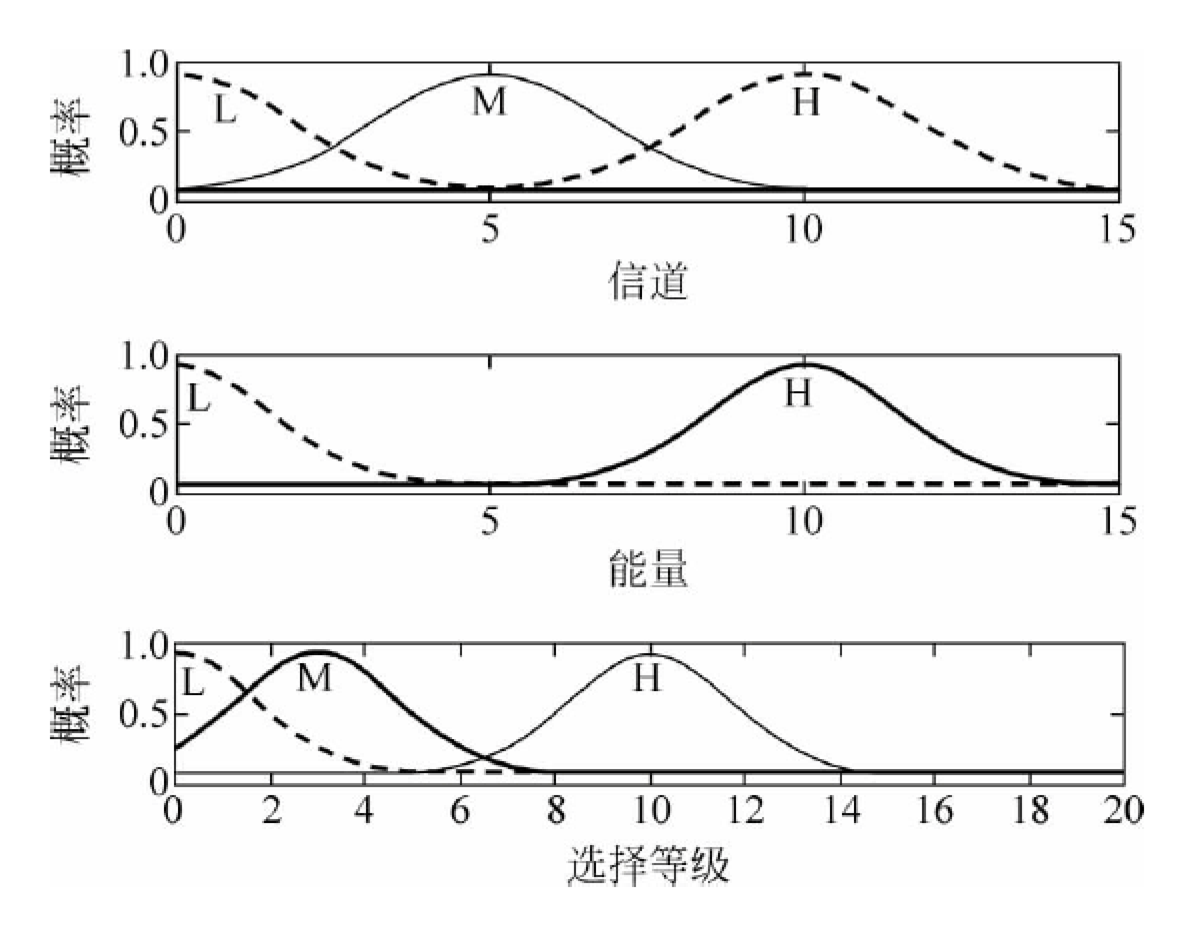


图 7-6 协调器节点选择的成员函数

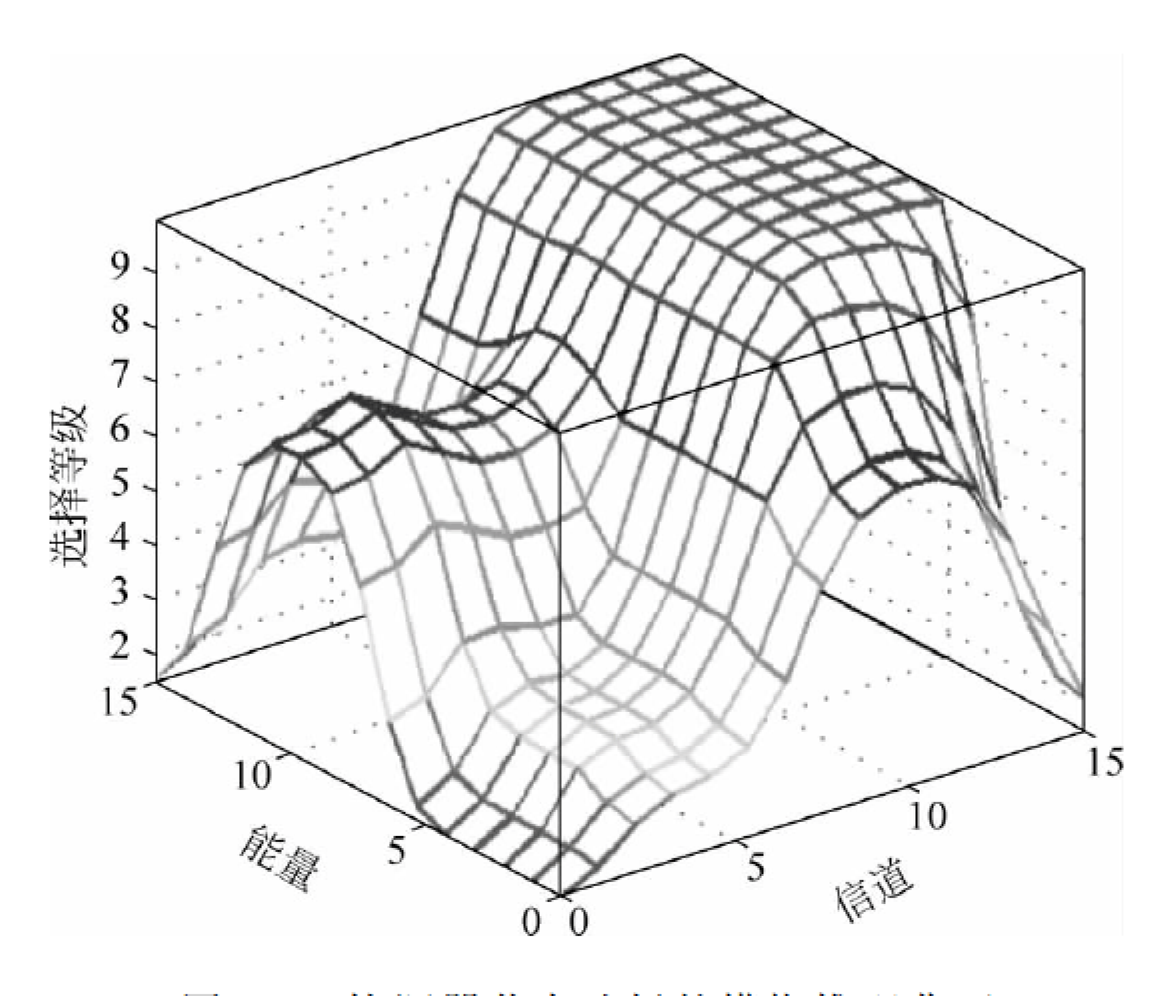


图 7-7 协调器节点选择的模糊推理曲面

图 7-8 显示了协调器节点能量对于协调器节点选择级别的影响，当 $E=6$ 时，协调器节点具有较高的能量，协调器节点的选择级别为最高，当 $E=3$ 时，协调器节点具有较低的能量，协调器节点的选择级别为最低。从图 7-9 可以看出，通过选择协调器节点，网络的吞吐量呈增长趋势。在具有 20 个节点的网络中，与随机的协调器节点选择方法相比，本章提出的博弈选择和模糊逻辑推理相结合的方法，其网络吞吐量达到 300b/s，而随机的协调器节点选择方法的网络吞吐量为 275b/s。

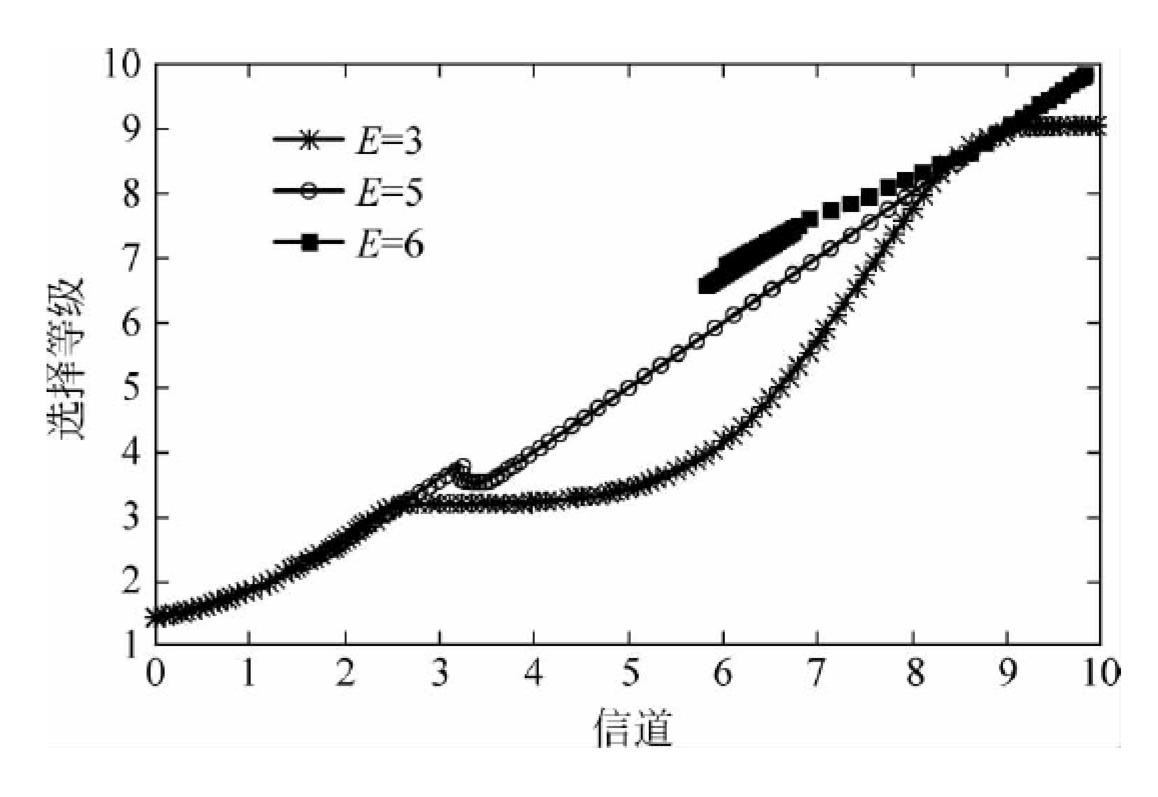


图 7-8　协调器节点的能量对协调器节点选择的影响

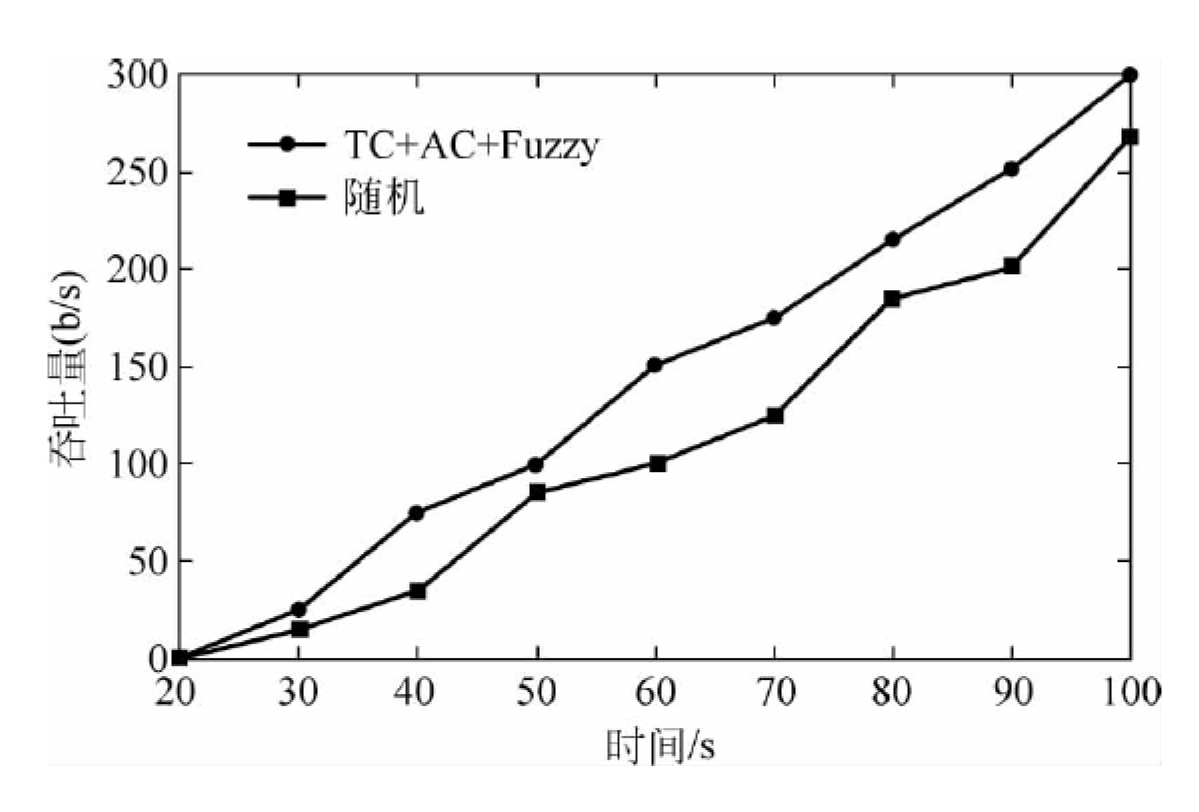


图 7-9　节点数为 20 时的网络吞吐量

7.6　小结

本章为受攻击的 ZigBee 无线传感器网络提出了一个主动防御机制，该机制保证了防御者能主动选择可靠的协调器节点来最小化网络的能量损失。使用演化博弈论模型化网络协调器的攻击防御问题为一个 2-player 的零和博弈。在协调器节点受到攻击时，为了提高网络性能和增加安全性，联盟成员节点需选择新的协调器节点。本章使用随机博弈论来响应协调器节点选择过程，把网络中的节点作为一个博弈参与者，来自于不同邻居节点的局部组合估计信息作为协调器节点选择的依据，使用模糊逻辑，估计协调器节点的可信状态，保证了被选择的协调器节点具有较高的安全性。实验结果表明，当协调器节点受攻击时，使用本章提出的算法能保证 ZigBee 无线传感器网络的吞吐量高于标准的 IEEE 802.15.4 协议。

第 8 章

面向传感云数据外包中心的信任演化机制研究

在面向传感云的数据外包中心认证系统中需要用户的个人信息和私有证书，由于传感云计算系统的动态性和证书私有性经常导致认证失败。为了解决这个问题，本章提出了基于动态证书博弈的认证系统框架。在证书认证博弈交互过程中，通过认证代理补偿一定的信任度来激励传感云用户出示更多的证书，以提高其信任度。传感云用户和认证协调器通过平衡证书泄露和信任补偿之间的关系来决定用户是否能够操作外包数据。这样认证协调器决定每次博弈信任度，认证代理决定信任度分配，整个动态证书博弈系统框架就可以模型化为一个 3 阶段博弈，最后，使用迭代博弈学习方法证明信任协同的稳定性。在传感云环境下，与传统的基于属性和本体的访问控制系统相比，本章提出的动态证书博弈机制提高了安全效用和认证性能。

8.1 引言

随着传感云计算技术的发展，为了降低计算资源的成本和提高数据使用的安全性，越来越多的企业把传感数据迁移到云中存储和管理。传感云数据外包中心部署在开放的环境中，服务提供者不仅需要向企业提供足够的软件、硬件和网络资源，而且还必须具有高效地创建、更新、访问外包传感数据的机制。同时，对于企业而言，通过把传感数据外包给传感云计算中心，企业可以集中处理其业务应用而无须部署软件和硬件来保存传感数据，这使得各种传感数据的存储和处理不再像传统的方式那样由企业雇用管理员来运行和维护，而是交由外包中心来完成。

然而，一旦外包中心中的传感数据泄露，将给企业造成巨大的损失。而且在开放计算环境中，企业的传感数据迁移到传感数据外包中心后，由于对传感云计算体系结构的安全由云服务提供者控制，真正的传感数据拥有者企业却对安全性的控制受到限制。鉴于此，传感云服务提供者必须为企业提供安全管理策略来提高数据的保密性和完整性。传感云服务提供者由于有太多的控制权，容易越过权限去修改企业的数据，这导致了传感云数据外包服务提供者和企业之间的较低信任关系，最终使得企业的传感数据面临巨大的安全挑战[390]。

用户和云平台之间的证书认证可保证数据的安全性。然而，由于传感云用户数量庞大、无线链路具有动态性，如何进行高效地认证是用户和传感云计算数据外包服务提供商均关

注的问题。这些问题一方面包括用户和传感云数据外包服务提供商之间的低信任关系；另一方面，由于目前缺乏信任的保证机制，容易造成外包数据被篡改、泄露。本质上，传感云中的外包数据保护不仅要对数据进行加密，还要实现用户和数据服务提供商之间的信任管理，通过加密和信任管理的双重保护，才能有效遏制数据的泄露。此外，目前的传感云终端网络环境开放，恶意软件泛滥，数据传输通道没有被很好保护。系统级的数据采集手段隐蔽，难以发现和管控，造成隐私失窃。用户的合法性在客户端验证潜藏着巨大的风险，把验证过程迁移到传感云数据外包中心可有效地降低安全风险。

本章将利用博弈论研究用户和传感云数据外包中心间的演化信任决策过程，从而揭示云计算网络中用户和传感云数据外包中心间的信任演化原理。根据博弈论的特点，本章将每个用户、认证代理和认证协调器看作博弈的参与者，将用户披露的证书看作博弈策略，然后根据各个用户能选择不同的证书披露策略的实际情况建立证书认证信任演化博弈模型及框架，并且为了研究证书泄露补偿机制对用户选择披露证书动作的影响，在证书认证信任演化博弈模型中整合用户证书披露程度、操作偏好泄露度和分配的信任度等参数。为了说明证书认证信任演化博弈模型的稳定性，通过信任协同学习动态方程探索证书认证信任演化博弈的演化稳定的最优信任状态。

本章的主要工作如下：

(1) 建立了适用于传感云数据外包中心的证书认证信任演化博弈模型来判断传感云用户的合法性。该模型考虑了证书披露、用户操作的敏感性和证书博弈过程中的信任度分配。

(2) 推导出了传感云用户证书披露的最优策略、认证代理信任度分配的最优策略、认证协调器的最优策略。

(3) 形式化了证书认证博弈的信任演化协同动态方程，使用 G-value 学习过程证明了证书认证信任演化博弈的稳定性。最后，讨论了混合证书认证策略。

本章其余章节安排如下：8.2 节介绍相关工作；8.3 节描述证书认证信任演化博弈模型；8.4 节阐述证书认证信任演化博弈的决策过程；8.5 节证明证书认证信任演化博弈的稳定性；8.6 节给出混合证书认证策略；8.7 节是仿真结果与分析；8.8 节给出本章小结。

本章涉及的符号含义如下：

$\widetilde{M}$表示认证协调器(AC)拥有证书的个数。

$\widetilde{N}$表示用户拥有证书的个数。

C^a 表示认证协调器的证书集合。

C^v 表示用户的证书集合。

S 表示用户集合。

AA 表示认证代理集合。

AC 表示认证协调器集合。

O 表示用户的操作集合。

O_k 表示用户的一个操作。

α_k 表示操作 O_k 的证书披露因子。

$n_{i,k}$表示用户 S_i 执行操作 O_k 披露证书的总数。

U_a 表示认证协调器的效用。

V_l 表示一个证书披露后，认证协调器 l 从认证代理获得的信任度。

n_l 表示认证协调器 l 对外包数据执行操作需要的证书数目。

K_l 表示认证协调器 l 发送证书给认证代理后的证书泄露度。

D_i 表示分配给用户的信任度。

U_b'表示认证代理的效用。

K_i 表示用户 S_i 执行操作时的证书披露度。

n_i 表示用户 S_i 披露证书的数目。

$C_{i,k}=0$ 表示用户 S_i 对于外包数据操作 O_k 决定不披露证书 C_i^v。

$C_{i,k}=1$ 表示用户 S_i 对于外包数据操作 O_k 决定披露证书 C_i^v。

β_k 表示操作偏好泄露因子。

$U_i^s(C_{i,k},C_{-i,k})$表示用户 S_i 的效用。

λ_i 表示用户 S_i 用于安全防御的资源损失率。

$C_{i,k}^*$表示用户 S_i 披露证书的最优策略。

$c(\lambda')$表示用户披露证书获得的收益函数。

λ'表示用户、认证协调器用于安全防御的资源损失率。

λ_l 表示认证协调器 l 用于安全防御的资源损失率。

U_b 表示 N 个用户认证时，认证代理的效用。

D_i^* 表示认证代理分配给用户 S_i 的最优信任度。

V_l^* 表示认证代理分配给认证协调器 l 的最优信任度。

K_l^* 表示认证协调器 l 的最优策略。

R 表示博弈参与者共享的信任状态空间。

$q_i(t)$表示在时刻 t 博弈参与者 i 用于安全防御的资源损失率。

$c_i(q_i(t))$表示在时刻 t 博弈参与者 i 的补偿信任度。

$c_i(\cdot)$表示用户、认证协调器披露证书的收益函数。

$u_i(t)$表示在时刻 t 博弈参与者 i 的博弈效用。

H 表示信任演化迭代轮号集合。

h 表示信任演化迭代轮号。

$c_i(q_{i,h}(t))$表示第 h 轮博弈的时刻 t 博弈参与者 i 的补偿信任度。

$u_{i,h}(t)$表示第 h 轮博弈的时刻 t 博弈参与者 i 的博弈效用。

$I_{\widetilde{c}}$ 表示博弈的最大轮号。

$\boldsymbol{B}_h(t)$表示证书的传递路径的邻接矩阵。

$a_{ijv,h}(t)$表示在第 h 轮博弈的时刻 t 存在一条证书传递路径 $S_i \rightarrow AA_j \leftarrow AC_v$。

$q_{ijv,h}(t)$表示第 h 轮博弈的时刻 t 证书传递路径 $S_i \rightarrow AA_j \leftarrow AC_v$ 用于安全防御的资源损失率。

$c_{ijv,h}(q_{ijv,h}(t))$表示第 h 轮博弈的时刻 t 证书传递路径 $S_i \rightarrow AA_j \leftarrow AC_v$ 的补偿信任度。

$q_{ij,h}(t)$表示第 h 轮博弈的时刻 t 证书传递路径 $S_i \rightarrow AA_j$ 用于安全防御的资源损失率。

$c_{ij,h}(q_{ij,h}(t))$表示第 h 轮博弈的时刻 t 证书传递路径 $S_i \rightarrow AA_j$ 的补偿信任度。

$q_{vj,h}(t)$表示第 h 轮博弈的时刻 t 证书传递路径 $AA_j \leftarrow AC_v$ 用于安全防御的资源损失率。

$c_{vj,h}(q_{vj,h}(t))$表示第 h 轮博弈的时刻 t 证书传递路径 $AA_j \leftarrow AC_v$ 的补偿信任度。

$T^r_{ijv}(t)$为博弈参与者 S_i、AA_j 和 AC_v 在时刻 t 期望的信任协同结果。

$e_{ijv,h}(t)$表示用户 S_i、认证代理 AA_j 和认证协调器 AC_v 的信任协同状态偏移误差。

$\mu_{ijv,h}(t)$表示第 h 轮博弈的时刻 t 证书传递路径 $S_i \rightarrow AA_j \leftarrow AC_v$ 的信任协同学习增益 G-value。

A 表示数据浏览权限证书。

B 表示数据增加权限证书。

C 表示数据删除权限证书。

D 表示数据修改权限证书。

E 表示数据迁移权限证书。

F 表示数据下载权限证书。

G 表示数据上传权限证书。

C^W 表示为不同的数据操作分配相应的信任等级权重。

W_A 表示数据浏览操作分配的信任权重。

W_B 表示数据插入操作分配的信任权重。

W_C 表示数据删除操作分配的信任权重。

W_D 表示数据修改操作分配的信任权重。

W_E 表示数据迁移操作分配的信任权重。

W_F 表示数据下载操作分配的信任权重。

W_G 表示数据上传操作分配的信任权重。

O_i 表示用户执行第 i 条策略组合操作获得的收益。

$I(\cdot)$表示指示函数。

$\eta_h(t)$表示用户在第 h 轮博弈中披露证书后获得的信任度。

$\tilde{c}$表示博弈轮数。

$\Phi_x(t)$表示用户在时刻 t 完成一个策略组合链上操作后被分配的信任度。

M_t 表示在时刻 t 用户可获得的活动策略链。

TD_i 表示从时刻 T_0 到 T_1，用户 S_i 获得的累积被分配的总信任度。

8.2　相关工作

防止恶意攻击者和传感云数据外包提供商窃取数据，为传感云用户提供一种安全、高效的数据保护措施尤为重要。相比于密码认证的受关注程度，目前仅有少量的文献关注动态证书认证，其中比较有代表性的工作有以下几个方面：

(1) 面向云计算安全方面。文献[391]分析并设计了面向云计算的基于 ADS(Authenticated Data Structures)的数据外包认证模型，扩展了数据一致性证据生成和验证算法。文献[392]阐述了在云计算环境中，需要使用证书来证明基础设施服务、虚拟服务器、用户、设备之间交互的合法性。它首先通过 PKI 产生需要的证书，然后扩展成包含有用户角色信息的 X.509 证书，经 Web 环境中的信任机构颁发证书。文献[393]针对类似于 iCloud 用户的密码泄露问题提出了基于云存储的密码管理机制(CSF-BPM)。随着移动互联网的普及，手机用户每

天访问云平台，但大多用户不善于对其用户名、密码和身份识别号管理，容易被攻击者偷窃，为了增强访问控制的安全级别，方便云用户身份识别和授权访问多个云服务提供者，文献[394]提出了第三方的身份识别和管理系统(IDMs)。面对云数据安全、滥用云服务、恶意的内部攻击，文献[395]提出了一个模型来识别不同的访问控制需求，对访问权限进行了控制。文献[396]对于云存储的数据保护提出了分布式的访问控制机制，能实现用户撤销、数据读取等权限。文献[397]为保证云用户上传和下载媒体数据的安全性，提出了数字水印的算法对用户和媒体服务提供者进行认证。文献[398]根据服务的信任感知来动态调度云服务。文献[399]针对个人浏览网络信息提出了基于博弈论的隐私保护框架。

(2) 面向外包数据中心数据保护方面。文献[400，401]针对外包数据的完整性，结合公钥加解密算法提出了一种高效的数据完整性审核机制。文献[402]通过分析向量和向量点积的代数性质，对外包数据进行正确性验证，有效地抵御了攻击威胁。文献[403]针对外包服务网络，根据已有的信息对缓存中的链接进行分析，使用域名服务雷达(DNS Radar)来探测恶意攻击。

(3) 面向无线传感器网络方面。文献[404]对车载通信网络，使用证书选择的方法对追踪的车辆进行隐私保护。由于车辆的移动性使得证书的个数受到限制，在解决这个问题时主要考虑了如何减少证书来快速地进行车辆的身份认证。在文献[405]中，对车载通信网络中的证书更新提供了一种分布式的分发机制，并使用批处理验证技术实现减少证书验证开销的目的。文献[406]针对无线传感器网络中的信任关系建立博弈模型，给出了无线传感器网络节点之间的信任演化动态方程，证明了信任演化的稳定性。文献[407-414]主要针对MANETs 网络提出了信任模型。

以上这些方法中，对于云计算中心数据的保护提出了验证和加密算法，但对于传感云资源管理的动态性所造成的动态认证问题，还未有相应的解决方案。在对车载网络通信中使用证书验证的方法对用户进行访问控制时，考虑了证书选取的随机性和快速认证的特征，但对于证书的泄露和信任度分配还未考虑。本章利用基于证书认证的信任演化博弈研究传感云用户和传感云数据外包中心动态认证的信任建立问题。通过建立传感云数据外包中心服务提供者和传感云用户间的博弈策略和模型来获得服务提供者与用户之间信任决策时的安全效用，引入激励机制促使服务提供者与传感云用户选择出示证书的策略。引入证书泄露机制、操作因子和学习增益使得服务提供者与传感云用户尽可能出示较少的证书获得较大信任度，并在较短的时间内完成认证。这一点有别于传统的基于属性和本体角色的认证方法。这些研究成果将为传感云数据外包中心服务提供者和传感云用户之间信任机制的设计提供理论基础。

8.3 证书认证信任演化博弈模型

8.3.1 传感云数据外包中心访问控制系统

传感云用户、传感云数据外包服务提供者之间的外包数据访问控制系统如图 8-1 所示。其中，整个访问控制系统由传感云用户、认证代理(服务访问点)、认证协调器组成。传感云数据外包中心通过云服务平台向用户提供访问控制服务。服务访问点作为认证代理负责对

用户和认证协调器出示的证书进行匹配。

假设认证协调器(AC)有$\widetilde{M}$个证书，用户有$\widetilde{N}$个证书。令$C^a=\{C_1^a,C_2^a,\cdots,C_{\widetilde{M}}^a\}$表示认证协调器的证书集合，$C^v=\{C_1^v,C_2^v,\cdots,C_{\widetilde{N}}^v\}$表示用户的证书集合，$S=\{S_1,S_2,\cdots,S_{\widetilde{s}}\}$表示用户集合，$AA=\{AA_1,AA_2,\cdots,AA_{\widetilde{b}}\}$表示认证代理集合，$AC=\{AC_1,AC_2,\cdots,AC_{\widetilde{a}}\}$表示认证协调器集合。所有的用户、认证代理、认证协调器通过理性博弈实现自身的效用最大化。

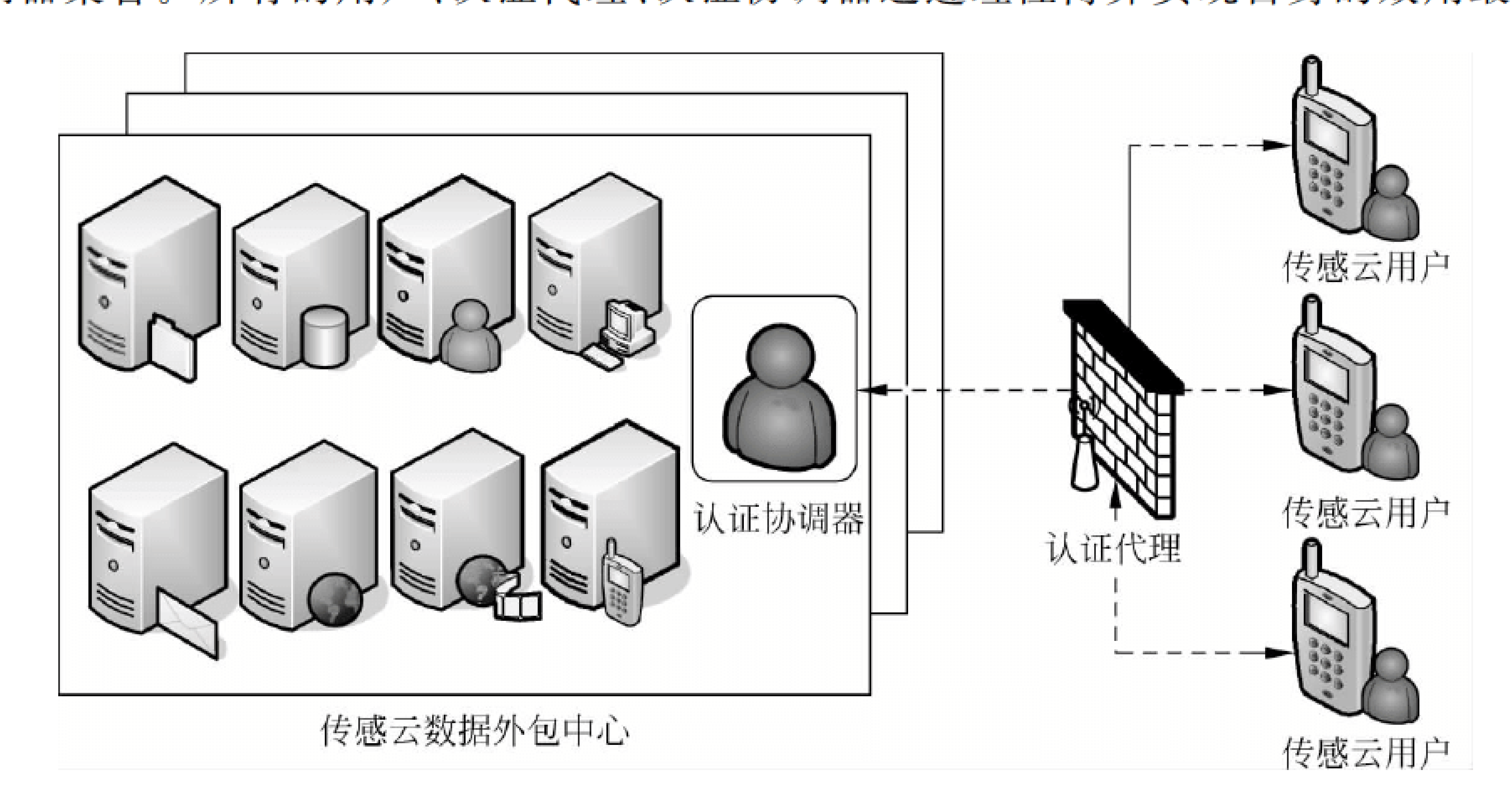


图 8-1　传感云数据外包中心访问控制系统

认证代理接收不同用户和认证协调器的证书并跟证书集C^a和C^v进行匹配。同时，认证代理计算用户和认证协调器每次证书披露后的信任度，对认证协调器每次的认证处理提高其认证等级。用户和认证协调器出示的证书是私有信息，这些信息被披露给第三方认证代理，认证代理分配给用户和认证协调器一定的信任度以补偿证书披露的损失，并且激励其披露更多的证书信息以提高信任度。用户和认证协调器根据证书私有敏感性、通过披露证书获得的信任度来决定是否继续披露证书。

8.3.2　私有证书披露敏感性

定义 8-1　证书认证信任演化博弈的信任度定义为信任演化博弈过程中为补偿私有证书披露而分配的一定数量的信任值。

用户披露不同种类的证书获得不同等级的数据访问权限。例如，披露数据迁移和修改证书表明用户有组合管理权限，但是，这容易造成高级的操作权限泄露。又如，当披露数据浏览证书时，仅仅表明需要检索数据，但是这会造成低级的操作权限泄露。记用户S_i的操作集合为$O=\{O_1,O_2,\cdots,O_a\}$，对操作O_k，$k\in\{1,2,\cdots,a\}$，相应的证书披露因子记为α_k。同时，披露证书造成的私有证书泄露与总的组合性证书披露数目成正比，这表明用户披露越少的证书，它将有越低的私有敏感性证书泄露。因此，把用户私有证书泄露度定义为$\alpha_k/n_{i,k}$，其中$n_{i,k}$表示用户S_i执行操作O_k披露证书的总数。

8.3.3　证书认证信任演化博弈的效用

用户使用披露的证书通过认证后，可以操作外包数据。对于用户的每次披露，令V_l表示一个证书披露后，认证协调器l从认证代理获得的信任度。认证协调器的效用定义为

$$U_a = V_l n_l - K_l n_l \tag{8-1}$$

式中，n_l 为认证协调器 l 对外包数据执行操作需要的证书数目；K_l 为认证协调器 l 发送证书给认证代理后的证书泄露度。

认证代理从认证协调器接收到外包数据操作所需的证书，再进行证书匹配通过后会把信任度 D_i 分配给用户 S_i 来补偿用户为了执行外包数据操作所披露的证书。认证代理(Broker)的效用定义为

$$U'_b = \sum_l K_l n_l + \sum_i K_i n_i - \sum_i D_i - \sum_l V_l \tag{8-2}$$

式中，K_i 为用户 S_i 执行操作时的证书披露度；n_i 为用户 S_i 披露证书的数目。

接下来，用户 S_i 通过博弈策略来决定是否披露证书，令 $C_{i,k}=0$ 表示用户 S_i 对于外包数据操作 O_k 决定不披露证书 C_i^v，$C_{i,k}=1$ 表示用户 S_i 对于外包数据操作 O_k 决定披露证书 C_i^v。每个用户为执行外包数据操作 O_k 出示证书后获得的信任度为

$$\sum_i D_i \frac{C_{i,k}}{\sum_i C_{i,k}} \tag{8-3}$$

而用户 S_i 执行外包数据操作的损失包括两方面：一方面是证书泄露；另一方面是操作偏好泄露。用户 S_i 执行数据操作 O_k 的证书和操作偏好泄露度表示为

$$\frac{(\alpha_k + \beta_k)}{\sum_i C_{i,k}} \tag{8-4}$$

式中，β_k 为操作偏好泄露因子。例如，用户通过 WiFi 接入传感云数据外包中心浏览传感数据，此时若无线链路被窃听者控制，虽然窃听者能截获证书，但窃听者还不能从截获的证书中获知此证书拥有的访问权限，于是，窃听者需进一步获得用户的操作行为偏好，根据用户的操作行为偏好，窃听者才可进一步推测出截获的证书是否具有浏览数据的权限。这里的 β_k 越大，表明用户泄露了越多的操作行为，窃听者就能获得越多与证书相关的操作权限信息，再根据这些信息假冒合法用户入侵传感云数据服务器来获取、篡改数据。用户 S_i 的效用可表示为

$$U_i^s(C_{i,k}, C_{-i,k}) = \sum_i D_i \frac{C_{i,k}}{\sum_i C_{i,k}} - \lambda_i \sum_i \frac{(\alpha_k + \beta_k) C_{i,k}}{\sum_i C_{i,k}} \tag{8-5}$$

式中，λ_i 为用户 S_i 用于安全防御的资源损失率。从式(8-5)中可看出，当 $\alpha_k+\beta_k$ 的值越小，且 D_i 的值越大，用户 S_i 的效用越大。也即，泄露较少的证书和操作行为偏好，并且获得了较高的信任度。同时，$\alpha_k+\beta_k$ 的值通过信任度的分配还影响了认证协调器、认证代理的效用。因此，证书认证信任演化博弈中各参与者之间的博弈过程实质就是增大用户、认证代理和认证协调器这三者之间效用的过程。也就是说，在博弈过程中通过适当地控制 $\alpha_k+\beta_k$ 的值来提高各自的安全效用，可以有效地防御攻击者对于用户的证书和操作行为偏好的窃取，从而提高用户访问传感云外包数据的安全性。

8.4 证书认证信任演化博弈

本章根据用户认证的交互过程，把证书认证信任演化博弈机制形式化描述为一个 3 阶段博弈。在第一阶段，用户首先向认证代理发送数据操作请求，认证代理根据其数据操作请

求向用户发送披露证书请求。为了最大化安全效用，认证代理和用户进行证书交互的同时，认证代理也与认证协调器进行证书交互。在第二阶段，用户首先披露其请求操作对应的证书，认证代理根据用户披露的证书分配信任值。在这个阶段，认证代理根据认证协调器和用户披露的证书来调整分配给认证协调器和用户的信任值。在第三阶段，认证协调器和用户根据认证代理分配的信任度决定是否继续披露证书。证书认证信任演化博弈中各参与者之间的交互过程如图 8-2 所示。

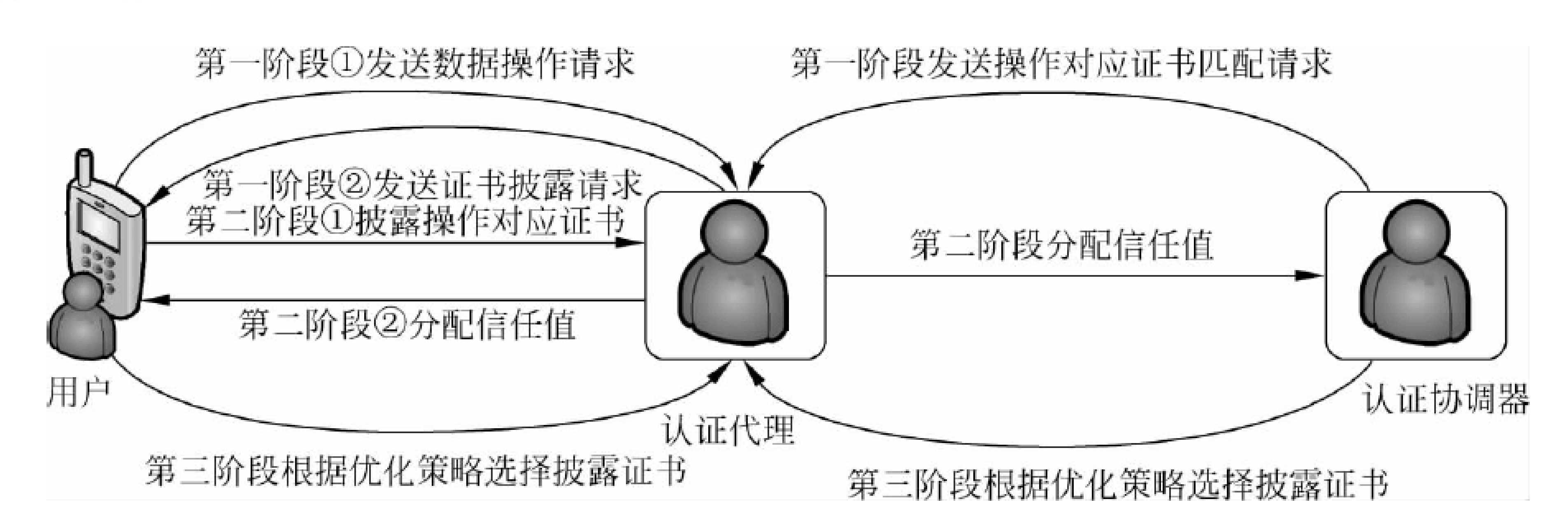


图 8-2　证书认证信任演化博弈交互过程

图 8-3 描述了证书认证信任演化博弈系统整体框架，它包含 7 个组件：用户、认证代理、认证协调器、信任状态识别、迭代学习增益 G 值、认证行动选择、证书披露策略。其中，用户、认证代理、认证协调器为博弈的参与者，通过这些参与者之间的协作认证使得系统信任度最终达到稳定状态。

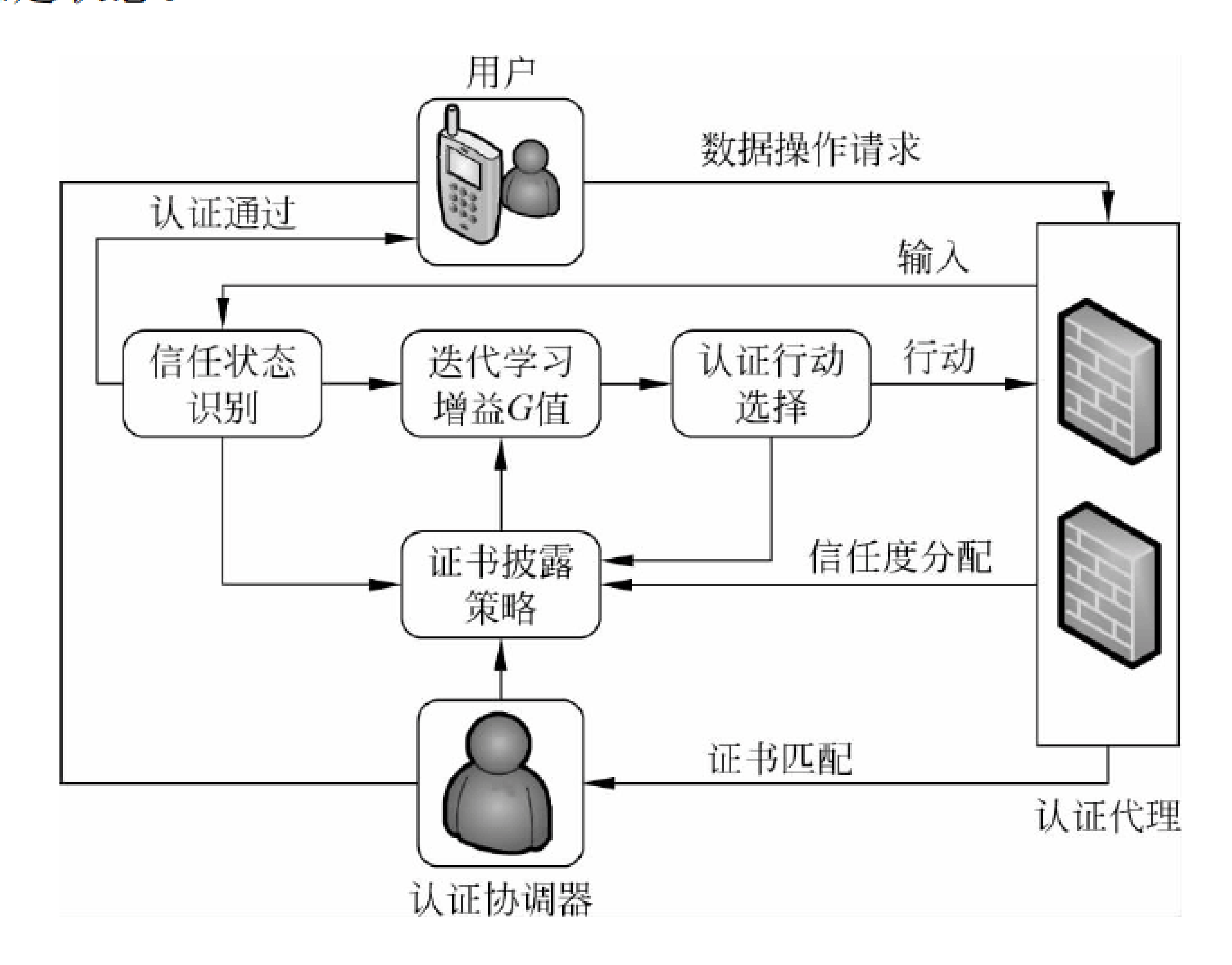


图 8-3　证书认证信任演化博弈系统框架

在图 8-3 中，信任状态识别指示了博弈参与者的信任度是否达到稳定，若系统框架中的博弈参与者的信任度达到稳定，则博弈参与者的安全效用达到最大，表明当前只需要泄露较少的证书和操作偏好，就能获得认证通过并进行数据操作。若系统框架中参与者的信任度还未达到稳定状态，此时，用户要不断地出示证书，然后认证协调器负责查找证书并提交给认证代理，认证代理再负责匹配证书。在此博弈过程中，用户、认证代理、认证协调器为了最

大化自己的效用，将选择博弈策略来达到信任协同的稳定状态。迭代学习增益 G 值是指用户、认证代理和认证协调器通过证书认证博弈获得累积效用的信任协同学习增益。认证行动选择是指终止认证或开始新一轮的认证。证书披露策略主要指用户将根据 $\alpha_k+\beta_k$ 和 D_i 的值来决定是否披露证书。

证书认证博弈的信任演化开始时，用户首先向认证代理发起数据操作请求，驱动认证代理输入系统的信任状态初始值，然后信任状态识别主要通过博弈状态方程描述博弈参与者之间信任度的演化状态。接下来，用户和认证协调器优化证书披露策略并选择行动。认证代理再根据用户和认证协调器的证书披露策略分配信任度。最后，用户、认证代理和认证协调器通过更新整个证书传递路径上的迭代学习增益 G 值使得整个信任演化系统框架达到纳什均衡。

8.4.1 用户披露证书的优化策略

对于同一个操作 O_k，每个用户都想最大化自己的安全效用，并出示较少的证书。因此，每个用户的安全效用不仅依赖于自身的决策，而且还依赖于其他用户的决策，用户 S_i 披露证书的优化策略可表示为

$$C_{i,k}^{*}=\arg\max_{C_{i,k}} U_i^s(C_{i,k},C_{-i,k}) \tag{8-6}$$

如果每个用户采用最优策略，则通过证书认证博弈，就可使得用户、认证代理、认证协调器之间的信任达到最优均衡。

定理 8-1 如果条件 $D_i>(\alpha_k+\beta_k)\lambda_i$ 成立，那么，在用户之间存在纳什均衡最优策略。

证明 由式(8-5)可以推导出

$$U_i^s(C_{i,k},C_{-i,k})=\sum_i(D_i-(\alpha_k+\beta_k)\lambda_i)\frac{C_{i,k}}{\sum_i C_{i,k}} \tag{8-7}$$

若 $D_i>(\alpha_k+\beta_k)\lambda_i$，则说明用户 S_i 披露证书，即 $C_{i,k}=1$，这样可增加 S_i 的安全效用 $U_i^s(C_{i,k},C_{-i,k})$。若 $D_i<(\alpha_k+\beta_k)\lambda_i$，则说明用户 S_i 不披露证书，即 $C_{i,k}=0$，这样用户 S_i 的证书泄露度较小，可避免降低安全效用 $U_i^s(C_{i,k},C_{-i,k})$。因此，S_i 通过决定 $C_{i,k}$ 的值能使用户保持最大安全效用 $U_i^s(C_{i,k},C_{-i,k})$，此时，$C_{i,k}^{*}$ 是用户 S_i 的最优策略。从而可以得出当 $D_i>(\alpha_k+\beta_k)\lambda_i$ 时，$U_i^s(C_{i,k},C_{-i,k})$ 保持非零状态，使得每个用户使用此种机制获得最优策略时，纳什均衡达到稳定状态。证毕。

8.4.2 认证代理信任演化博弈策略

当用户 S_i 根据定理 8-1 作出决策后，认证代理感知到用户 S_i 策略的变化，再决定分配给用户 S_i 的信任度 D_i，从而认证代理能决策其最优的信任度分配策略。

定理 8-2 若有 N 个用户同时发起认证请求时，认证代理存在最优信任度分配策略使其效用最大化。

证明 设用户披露证书获得的收益函数为

$$c(\lambda')=1-\mathrm{e}^{-c_m\lambda'} \tag{8-8}$$

式中，c_m 为常量；λ' 为用户、认证协调器用于安全防御的资源损失率，其中资源包括执行外包数据操作泄露的证书和操作偏好资源。当传感云数据外包中心同时有 N 个用户发出证

书认证请求时，用户 S_i 用于安全防御的资源损失率为

$$\lambda_i = \frac{D_i}{\alpha_k + \beta_k} \tag{8-9}$$

用户 S_i 披露证书的数目为

$$n_i = N \int_0^{\frac{D_i}{\alpha_k + \beta_k}} c(\lambda_i) \mathrm{d}\lambda_i \tag{8-10}$$

当传感云数据外包中心同时也有 M 个认证协调器发出证书匹配请求时，认证协调器 l 用于安全防御的资源损失率为

$$\lambda_l = \frac{V_l}{\alpha_k + \beta_k} \tag{8-11}$$

认证协调器 l 对外包数据执行操作需要的证书数目为

$$n_l = M \int_0^{\frac{V_l}{\alpha_k + \beta_k}} c(\lambda_l) \mathrm{d}\lambda_l \tag{8-12}$$

由式(8-2)可以推导出 N 个用户认证时，认证代理的效用为

$$U_b = \sum_i \left(K_i N \int_0^{\frac{D_i}{\alpha_k + \beta_k}} c(\lambda_i) \mathrm{d}\lambda_i - D_i \right) + \sum_l \left(K_l M \int_0^{\frac{V_l}{\alpha_k + \beta_k}} c(\lambda_l) \mathrm{d}\lambda_l - V_l \right) \tag{8-13}$$

对 U_b 分别求 D_i 和 V_l 的一阶偏导得

$$\frac{\mathrm{d}U_b}{\mathrm{d}D_i} = \sum_i \left[\frac{K_i N}{\alpha_k + \beta_k} \cdot c\left(\frac{D_i}{\alpha_k + \beta_k} \right) - 1 \right] \tag{8-14}$$

$$\frac{\mathrm{d}U_b}{\mathrm{d}V_l} = \sum_l \left[\frac{K_l M}{\alpha_k + \beta_k} \cdot c\left(\frac{V_l}{\alpha_k + \beta_k} \right) - 1 \right] \tag{8-15}$$

当 $\frac{\mathrm{d}U_b}{\mathrm{d}D_i}=0$、$\frac{\mathrm{d}U_b}{\mathrm{d}V_l}=0$ 时，认证代理获得最优信任度分配策略解，其中，认证代理分配给用户 S_i 的最优信任度为

$$D_i^* = (\alpha_k + \beta_k) c^{-1} \left(\frac{\alpha_k + \beta_k}{K_i N} \right) \tag{8-16}$$

认证代理分配给认证协调器 l 的最优信任度为

$$V_l^* = (\alpha_k + \beta_k) c^{-1} \left(\frac{\alpha_k + \beta_k}{K_l M} \right) \tag{8-17}$$

令

$$c^{-1} \left(\frac{\alpha_k + \beta_k}{K_i N} \right) = X \tag{8-18}$$

则

$$c(X) = \frac{\alpha_k + \beta_k}{K_i N} \tag{8-19}$$

由式(8-8)得

$$X = -\frac{1}{c_m} \ln \left(1 - \frac{\alpha_k + \beta_k}{K_i N} \right) \tag{8-20}$$

代入式(8-16)和式(8-17)中，可分别得到

$$D_i^* = (\alpha_k + \beta_k) \left[-\frac{1}{c_m} \ln \left(1 - \frac{\alpha_k + \beta_k}{K_i N} \right) \right] \tag{8-21}$$

$$V_l^* = (\alpha_k + \beta_k)\left[-\frac{1}{c_m}\ln\left(1 - \frac{\alpha_k + \beta_k}{K_l M}\right)\right] \tag{8-22}$$

证毕。

8.4.3 认证协调器信任演化博弈策略

认证协调器通过观察可感知认证代理的策略变化，从而选择其最优策略。

定理 8-3 若 U_a 连续且有界，且满足 $\frac{d(\ln U_a)}{dK_l} < 0$ 时，则认证协调器存在最优策略解 K_l^*，使其效用最大。

证明 由式(8-1)可推导出

$$U_a = (V_l - K_l)n_l = (V_l - K_l)M\int_0^{\frac{V_l}{\alpha_k+\beta_k}} c(\lambda_l)d\lambda_l \tag{8-23}$$

可得 U_a 连续且 $U_a \leqslant V_l$，所以，把 U_a 取对数后再求 K_l 的一阶偏导得

$$\frac{d(\ln U_a)}{dK_l} = -\frac{1}{V_l - K_l} + \frac{c\left(\frac{V_l^*}{\alpha_k + \beta_k}\right)}{\int_0^{\frac{V_l^*}{\alpha_k+\beta_k}} c(\lambda_l)d\lambda_l} \cdot \frac{d\left(\frac{V_l^*}{\alpha_k + \beta_k}\right)}{dK_l} \tag{8-24}$$

由式(8-8)得

$$c\left(\frac{V_l^*}{\alpha_k + \beta_k}\right) = 1 - e^{-c_m\left(\frac{V_l^*}{\alpha_k+\beta_k}\right)} \tag{8-25}$$

由式(8-22)，可得

$$\frac{d\left(\frac{V_l^*}{\alpha_k + \beta_k}\right)}{dK_l} = \frac{1}{\alpha_k + \beta_k} \cdot \frac{d(V_l^*)}{dK_i} = -\frac{\alpha_k + \beta_k}{c_m K_l(K_l M - (\alpha_k + \beta_k))} \tag{8-26}$$

把式(8-25)和式(8-26)代入式(8-24)可得

$$\frac{d(\ln U_a)}{dK_l} = -\frac{1}{V_l - K_l} - \frac{1 - e^{-c_m\left(\frac{V_l^*}{\alpha_k+\beta_k}\right)}}{\int_0^{\frac{V_l^*}{\alpha_k+\beta_k}} c(\lambda_l)d\lambda_l} \cdot \frac{\alpha_k + \beta_k}{c_m K_l(K_l M - (\alpha_k + \beta_k))} \tag{8-27}$$

由于 $\frac{d(\ln U_a)}{dK_l} < 0$，故 $\frac{d(\ln U_a)}{dK_l}$ 是关于自变量 K_l 的递减函数。所以，当 $\frac{d(\ln U_a)}{dK_l} = 0$ 时，存在优化解 K_l^*，即认证协调器的最优策略，使其效用最大。证毕。

8.5 证书认证信任演化博弈的稳定性分析

信任度的自适应学习过程对于每个博弈参与者来说是相同的。在用户、认证代理、认证协调器三者中的信任度学习的结果是使得它们达到信任协同，也即保持较高的信任度聚集效应。假设 $\tilde{m}$ 个由用户、认证代理、认证协调器组成的动态证书认证访问控制网络，共享相同的信任状态空间 R。令 t 为一个时间变量，每个参与者在时刻 t 的信任演化博弈状态方程为

$$c_i(q_i(t+1)) = c_i(q_i(t)) + u_i(t), \quad i = 1, \cdots, \tilde{m} \tag{8-28}$$

式中，$q_i(t)$ 为在时刻 t 博弈参与者 i 用于安全防御的资源损失率；$c_i(q_i(t))$ 为在时刻 t 博弈

参与者 i 的补偿信任度；$u_i(t)$为在时刻 t 博弈参与者 i 的博弈效用。令 $H=\{I_1,I_2,\cdots,I_{\tilde{\tau}}\}$ 表示信任演化迭代轮号集合，$h\in H$ 表示信任演化迭代轮号。在第 h 轮博弈，式(8-28)中博弈参与者 i 的补偿信任度和博弈效用分别表示为 $c_i(q_{i,h}(t))$和 $u_{i,h}(t)$，第 i 个参与者的补偿信任度动态演化方程表示为

$$c_i(q_{i,h}(t+1))=c_i(q_{i,h}(t))+u_{i,h}(t),\quad i=1,\cdots,\tilde{m} \tag{8-29}$$

式(8-29)表示了在证书认证信任演化博弈的每一轮中，用户 S_i、认证代理 AA_j 和认证协调器 AC_v 的博弈策略是动态变化的，从而引起它们的信任度也在不断变化。若证书的传递路径为 $S_i\rightarrow AA_j\leftarrow AC_v$，则表示它对应于第 h 轮博弈的时刻 t，从 S_i 到 AA_j 和 AC_v 到 AA_j 的一条认证路径。把证书的传递路径的邻接矩阵表示为

$$\boldsymbol{B}_h(t)=[a_{ijv,h}(t)] \tag{8-30}$$

式中，$a_{ijv,h}(t)$为在第 h 轮博弈的时刻 t 存在一条证书传递路径 $S_i\rightarrow AA_j\leftarrow AC_v$。

对任意 3 个博弈参与者 S_i、AA_j 和 AC_v，把

$$c_{ijv,h}(q_{ijv,h}(t))=c_{ij,h}(q_{ij,h}(t))-c_{vj,h}(q_{vj,h}(t)) \tag{8-31}$$

作为它们的信任协同演化方程。其中，$c_{ijv,h}(q_{ijv,h}(t))$表示第 h 轮博弈的时刻 t 证书传递路径 $S_i\rightarrow AA_j\leftarrow AC_v$ 的补偿信任度，$c_{ij,h}(q_{ij,h}(t))$表示第 h 轮博弈的时刻 t 证书传递路径 $S_i\rightarrow AA_j$ 的补偿信任度，$c_{vj,h}(q_{vj,h}(t))$表示第 h 轮博弈的时刻 t 证书传递路径 $AA_j\leftarrow AC_v$ 的补偿信任度。令 $T^r_{ijv}(t)$为博弈参与者 S_i、AA_j 和 AC_v 在时刻 t 期望的信任协同结果，从而可把证书认证的信任演化是否稳定的问题转化为用户 S_i、认证代理 AA_j 和认证协调器 AC_v 之间的信任协同问题，用户、认证代理和认证协调器的信任协同状态偏移误差为

$$e_{ijv,h}(t)=T^r_{ijv}(t)-c_{ijv,h}(q_{ijv,h}(t)) \tag{8-32}$$

根据证书披露信息和认证代理的信任值分配信息，形式化信任协同效用为

$$u_{i,h+1}(t)=u_{i,h}(t)+\sum_{i\in S,j\in AA,v\in AC,h\in H}\mu_{ijv,h}(t)a_{ijv,h}(t)\times[T^r_{ijv}(t)-c_{ijv,h}(q_{ijv,h}(t))] \tag{8-33}$$

其中，

$$\mu_{ijv,h}(t)=\begin{cases}\mu_i\in(0,1], & 若\ C_{i,k}=1\\ 0, & 其他\end{cases} \tag{8-34}$$

表示第 h 轮博弈的时刻 t 证书传递路径 $S_i\rightarrow AA_j\leftarrow AC_v$ 的信任协同学习增益 G 值。这样可得到博弈参与者后一轮与前一轮的效用增量为

$$\Delta u=u_{i,h+1}(t)-u_{i,h}(t)=\sum_{i\in S,j\in AA,v\in AC,h\in H}\mu_{ijv,h}(t)a_{ijv,h}(t)\times[T^r_{ijv}(t)-c_{ijv,h}(q_{ijv,h}(t))] \tag{8-35}$$

若 $T^r_{ijv}(t)-c_{ijv,h}(q_{ijv,h}(t))=0$，则 $\Delta u=0$，说明证书认证的演化博弈效用不再增长，演化达到稳定状态，此时的 $c_{ijv,h}(q^*_{ijv,h}(t))$为最优补偿信任度，用户、认证代理和认证协调器不需要再披露证书获得信任度而提高自己的安全效用。

8.6 混合证书认证策略

在传感云数据外包中心，用户操作数据的证书集表示为

$$C_M=\{A,B,C,D,E,F,G\} \tag{8-36}$$

式中，A 为数据浏览权限证书；B 为数据增加权限证书；C 为数据删除权限证书；D 为数据修改权限证书；E 为数据迁移权限证书；F 为数据下载权限证书；G 为数据上传权限证书。当一个证书博弈过程中只包含针对一种数据操作的证书策略，叫做单证书信任认证策略。当一个证书博弈过程中包含针对多种数据操作的证书策略，叫做混合证书信任认证策略。实际上，当用户对外包数据进行组合操作时，需要混合证书信任认证策略。即，对外包数据进行组合操作时，获得不同数据操作权限要出示不同的证书，混合安全策略如表 8-1 所示。

表 8-1　混合安全策略

策略	浏览	增加	删除	修改	迁移	下载	上传
P_1	*		*				
P_2		*		*			
P_3	*	*	*	*			
P_4					*	*	*
P_5		*	*	*	*	*	*
P_6	*	*		*	*	*	*
P_7	*	*	*	*	*	*	

在表 8-1 中，P_i，$i\in\{1,2,\cdots,7\}$，表示数据操作策略，* 表示组合策略具有的权限。不同的数据操作对应的信任等级权重为

$$C^W=\{1,2,3,4,5,6,7\} \tag{8-37}$$

权重越大信任级别越高，证书的保护程度越高。令数据操作和证书信任等级 L 的权重 W 表示为

$$C_L^W=\{A:2,B:3,C:7,D:6,E:5,F:4,G:1\} \tag{8-38}$$

其中，数据操作 C 对应的证书信任等级最高，其次是数据操作 D。令 W_A 表示数据浏览操作分配的信任权重，W_B 表示数据插入操作分配的信任权重，W_C 表示数据删除操作分配的信任权重，W_D 表示数据修改操作分配的信任权重，W_E 表示数据迁移操作分配的信任权重，W_F 表示数据下载操作分配的信任权重，W_G 表示数据上传操作分配的信任权重。用户执行第 i 条策略组合操作获得的收益可表示为

$$O_i=W_A^iI_A^i+W_B^iI_B^i+W_C^iI_C^i+W_D^iI_D^i+W_E^iI_E^i+W_F^iI_F^i+W_G^iI_G^i \tag{8-39}$$

式中，$I(\cdot)$为指示函数，当用户出示数据操作对应的证书时，函数值为 1；否则为 0。例如，在表 8-1 中，对于认证代理中的 P_6 策略，若分配的操作信任权重分别为：$W_A=2$，$W_B=3$，$W_C=0$，$W_D=6$，$W_E=5$，$W_F=4$，$W_G=1$，且各指示函数的值分别为：$I_A=1$，$I_B=0$，$I_C=0$，$I_D=1$，$I_E=0$，$I_F=0$，$I_G=1$，则用户获得收益为 $O_6=9$。用户执行第 i 条策略组合操作获得的收益由认证代理来计算，同时认证代理还实现了组合策略的层次结构。根据表 8-1，在图 8-4 中描述了进行证书和操作信任匹配的投影层次结构，在结构图中按照不同的策略可分为多个层次，通过这个层次结构进行证书和操作信任的匹配，其中，$\{P_1,P_2,P_3,P_4,P_5,P_6,P_7\}$构成策略组合链。$\{A,B,C,D,E,F,G\}$构成证书和操作信任链，它们组成了混合证书认证策略。混合证书认证策略的信任度分配包括两次信任度分配，一次是在每轮披露证书过程中分配信任度；另一次是完成策略组合链上的所有操作后再次分配信任度。

用户 S_i 多轮披露证书后累积被分配的信任度表示为

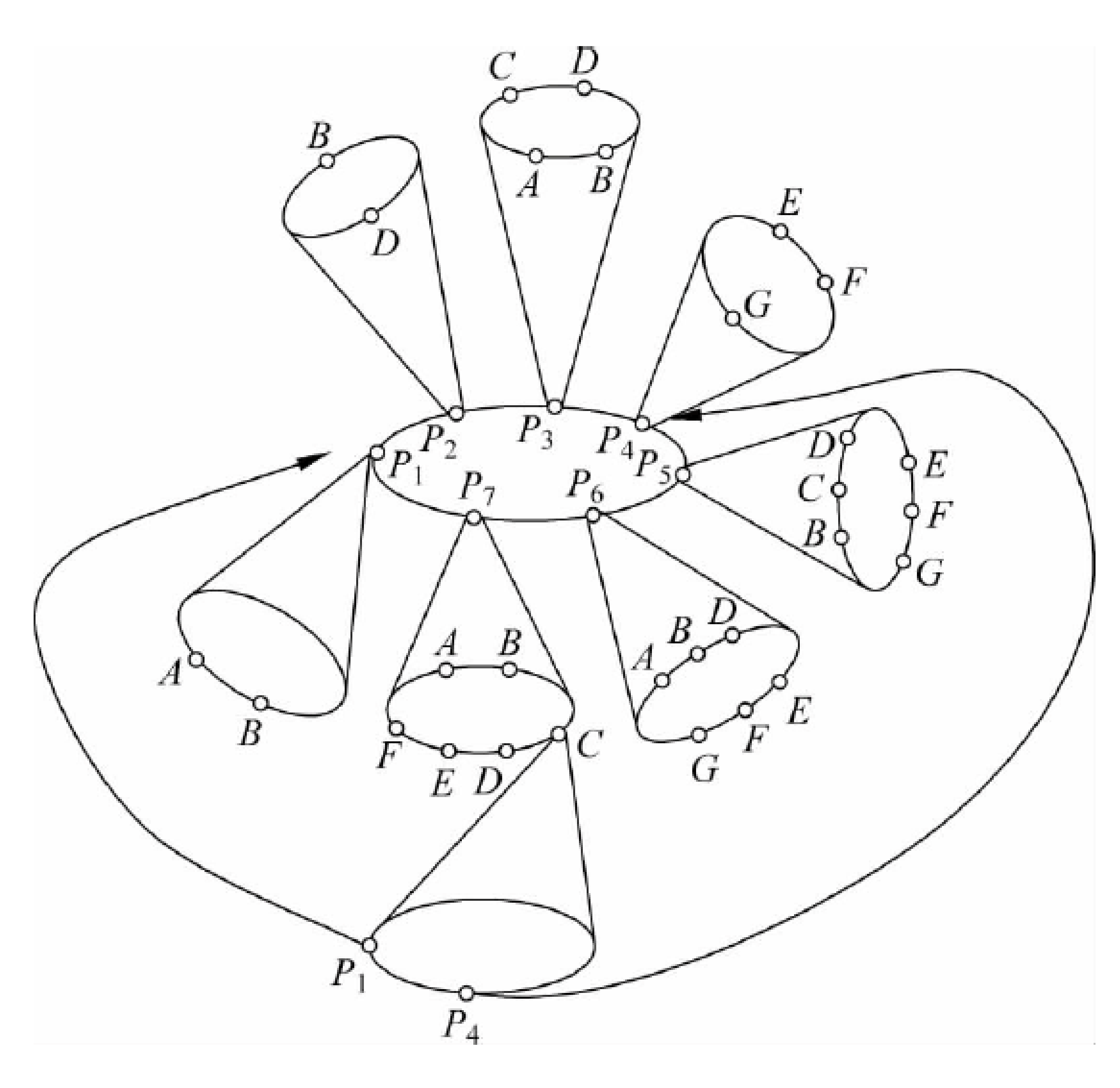


图 8-4　信任策略链和证书链投影层次结构

$$\eta_i(t) = \sum_{h=1}^{\tilde{c}} \eta_h(t) \tag{8-40}$$

式中，$\eta_h(t)$为用户在第 h 轮博弈中披露证书后获得的信任度；$\tilde{c}$为博弈轮数。用户 S_i 完成所有策略组合链上操作后累积被分配的信任度为

$$\Phi_i(t) = \sum_{x=1}^{M_i} \Phi_x(t) \tag{8-41}$$

式中，$\Phi_x(t)$为用户在时刻 t 完成一个策略组合链上操作后被分配的信任度；M_t 为在时刻 t 用户可获得的活动策略链。认证代理和认证协调器可使用这些策略链控制用户访问操作的信任权限。

式(8-40)和式(8-41)是以时间为自变量的函数，随着时间的推移，用户通过披露不同的证书链来获得不同的策略组合链，从而执行不同的外包数据操作。从时刻 T_0 到 T_1，用户 S_i 获得的累积被分配的总信任度为

$$TD_i = \int_{T_0}^{T_1} (\eta_i(t) + \Phi_i(t))\mathrm{d}t \tag{8-42}$$

8.7　实验

在这部分对博弈过程进行仿真并且评估其性能。在仿真中考虑 3 个方面问题：一是用户证书和操作偏好的泄露对于用户信任度的影响；二是用户证书和操作偏好的泄露对其效用的影响；三是确定认证协调器终止一轮认证过程的合适参数和信任协同学习增益对演化稳定状态的影响。

(1) 用户、认证协调器选择最优策略对其安全效用的影响。

图 8-5 给出了随着用户选择最优策略 $C_{i,k}^*$ 披露证书，使得每次证书博弈中用户的平均信任度随着其证书披露获得收益而增长，从而最大化用户的安全效用 $U_i^s(C_{i,k}, C_{-i,k})$。图 8-6

给出了在博弈过程中，认证协调器选择最优策略 K_i^*，使得认证协调器的效用随着证书披露而增长。值得说明的是，认证代理获得的效用越大，会使认证代理分配更多的信任度来激励用户披露证书。

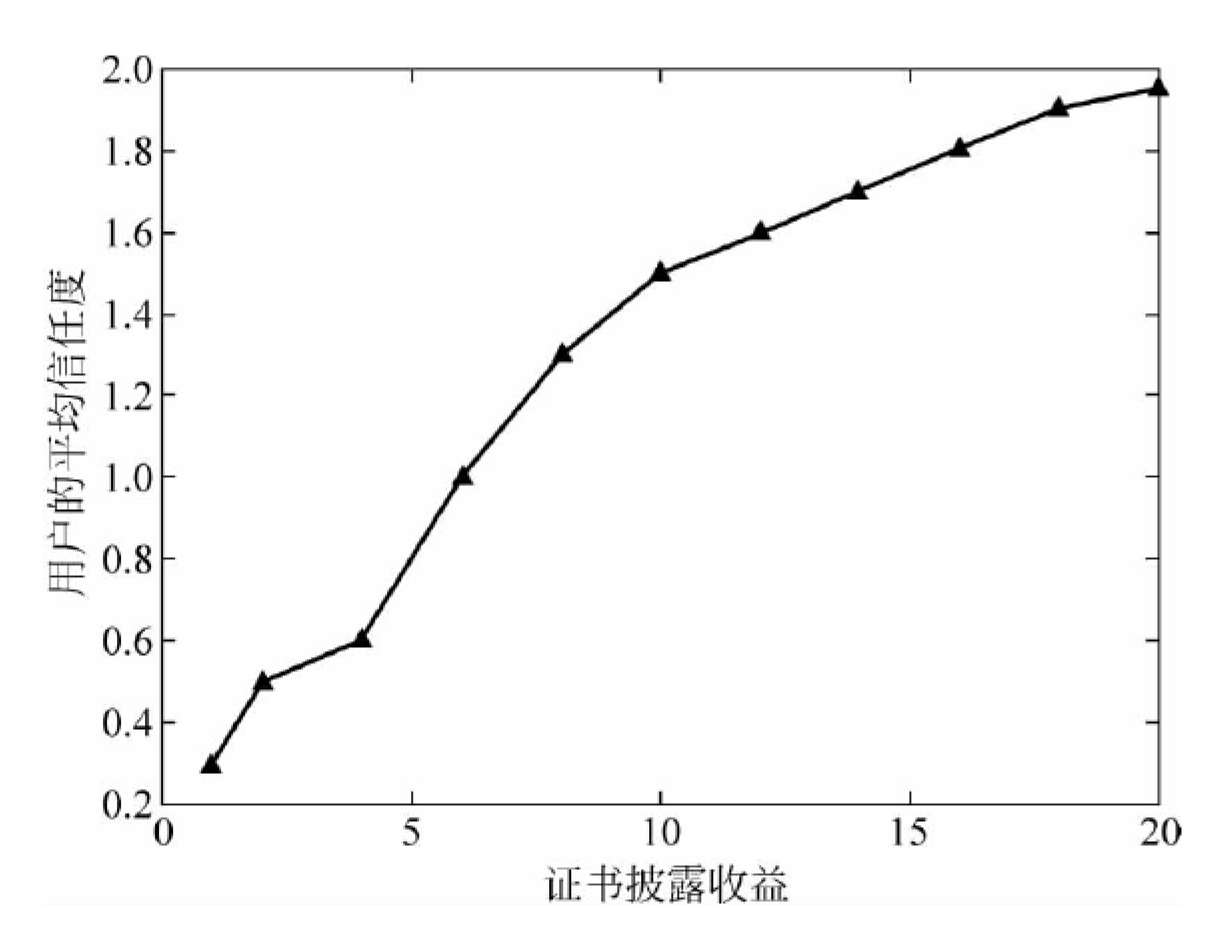


图 8-5 用户每次博弈获得的信任度

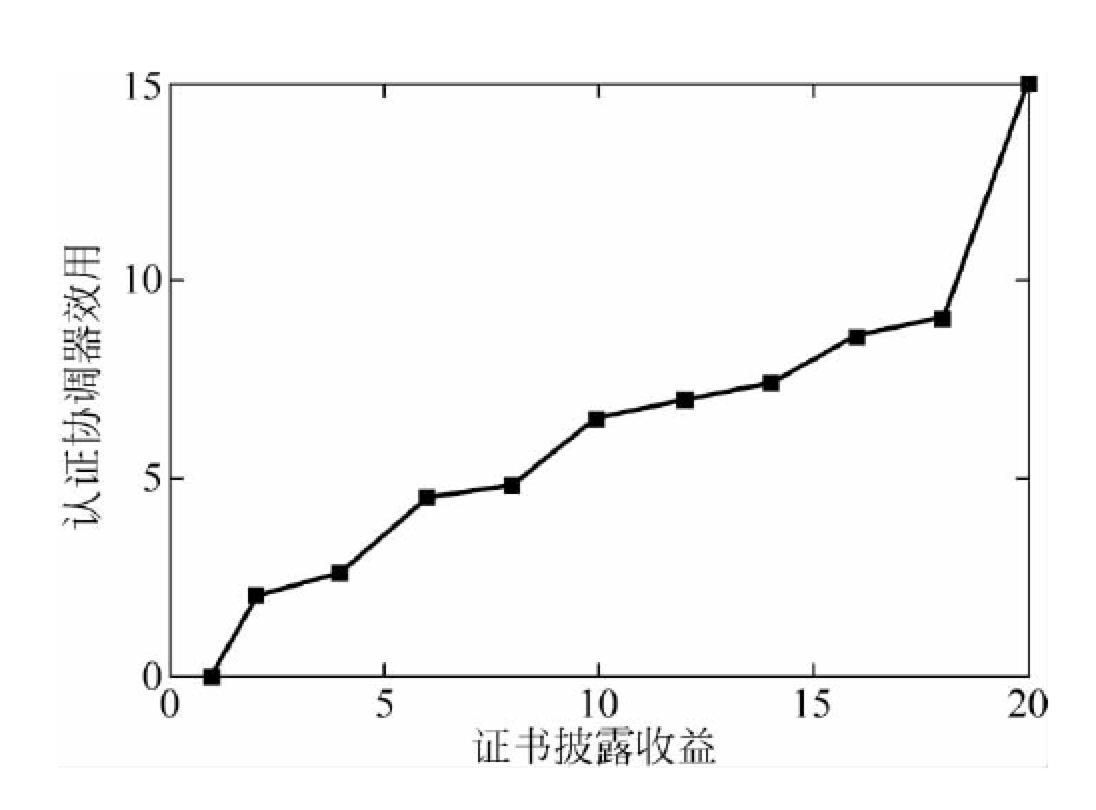


图 8-6 认证协调器证书博弈获得的效用

(2) 用户证书披露因子和操作偏好对其信任度的影响。

当 $D_i > (\alpha_k + \beta_k)\lambda_i$ 时，用户开始披露证书来获取信任度。α_k、β_k 的取值影响用户是否披露证书及获得信任度的大小。在图 8-7 中，当用户证书披露因子和操作偏好增长时，导致很高的证书和操作偏好泄露度。用户的证书披露和操作偏好泄露越多，其获得的信任度就越高，当 $\alpha_k = \beta_k = 3.5$ 时，用户的信任度达到最大值约为 430。通过归一化处理后，信任度经过认证代理分配给用户。

(3) 多轮证书披露累积信任度和所有策略组合链上操作完成后的信任度分配。

总的信任度分配量由式(8-42)中 $\eta_i(t)$ 和 $\Phi_i(t)$ 值决定，即多轮证书披露累积信任度和所有策略组合链上操作完成后的信任度分配。在图 8-8 中，认证代理分配信任度补偿用户证书和操作偏好泄露，使得用户的信任度增加。在仿真中，使用 η 表示 $\eta_i(t)$，ϕ 表示 $\Phi_i(t)$，随着 η 和 ϕ 值的增加，认证代理分配信任度增大，当 $\eta = \phi = 3.5$ 时，分配的信任度可达 8.2，经过归一化函数处理后分配给用户。

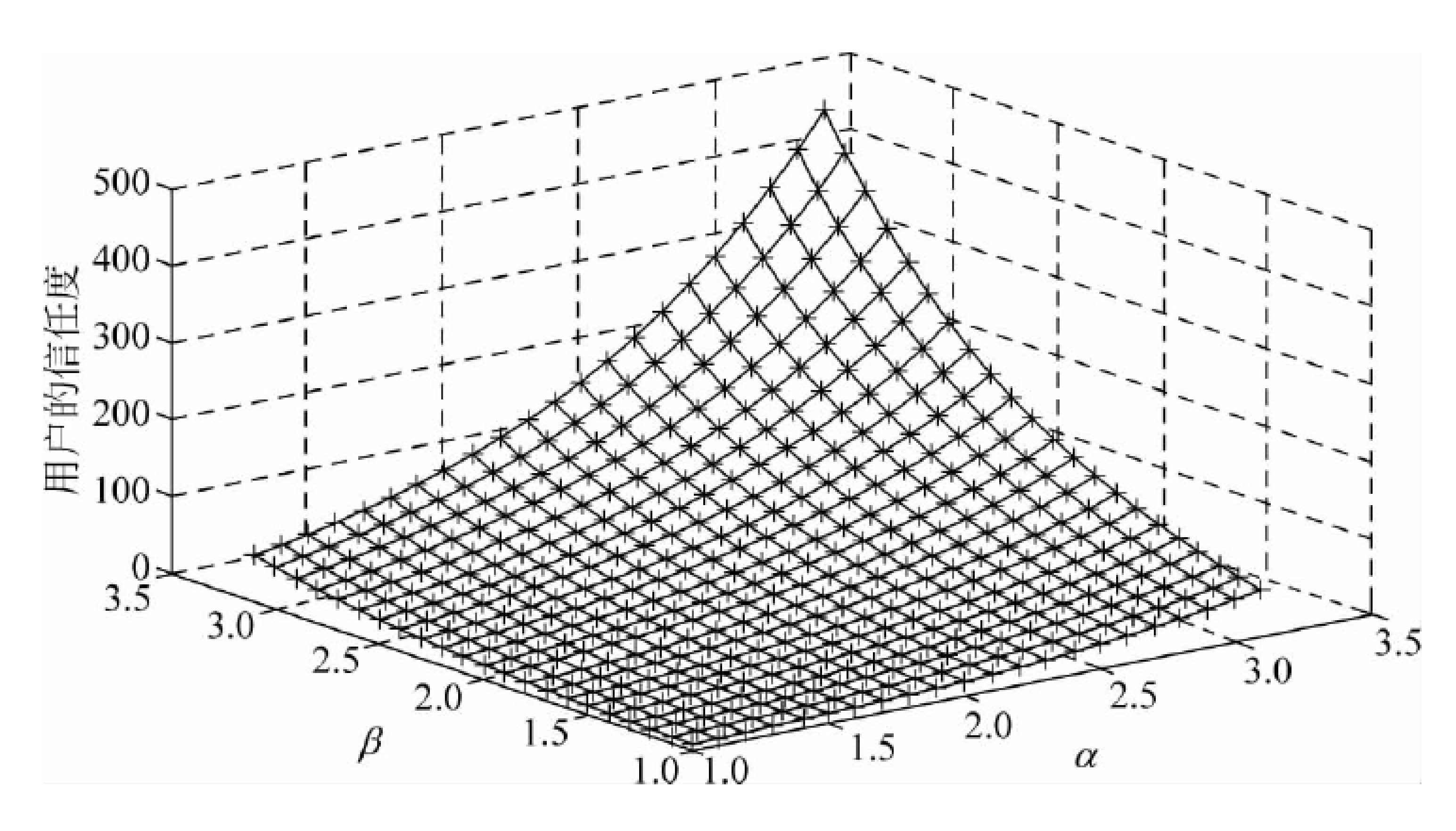


图 8-7　用户证书披露因子和操作偏好对其信任度的影响

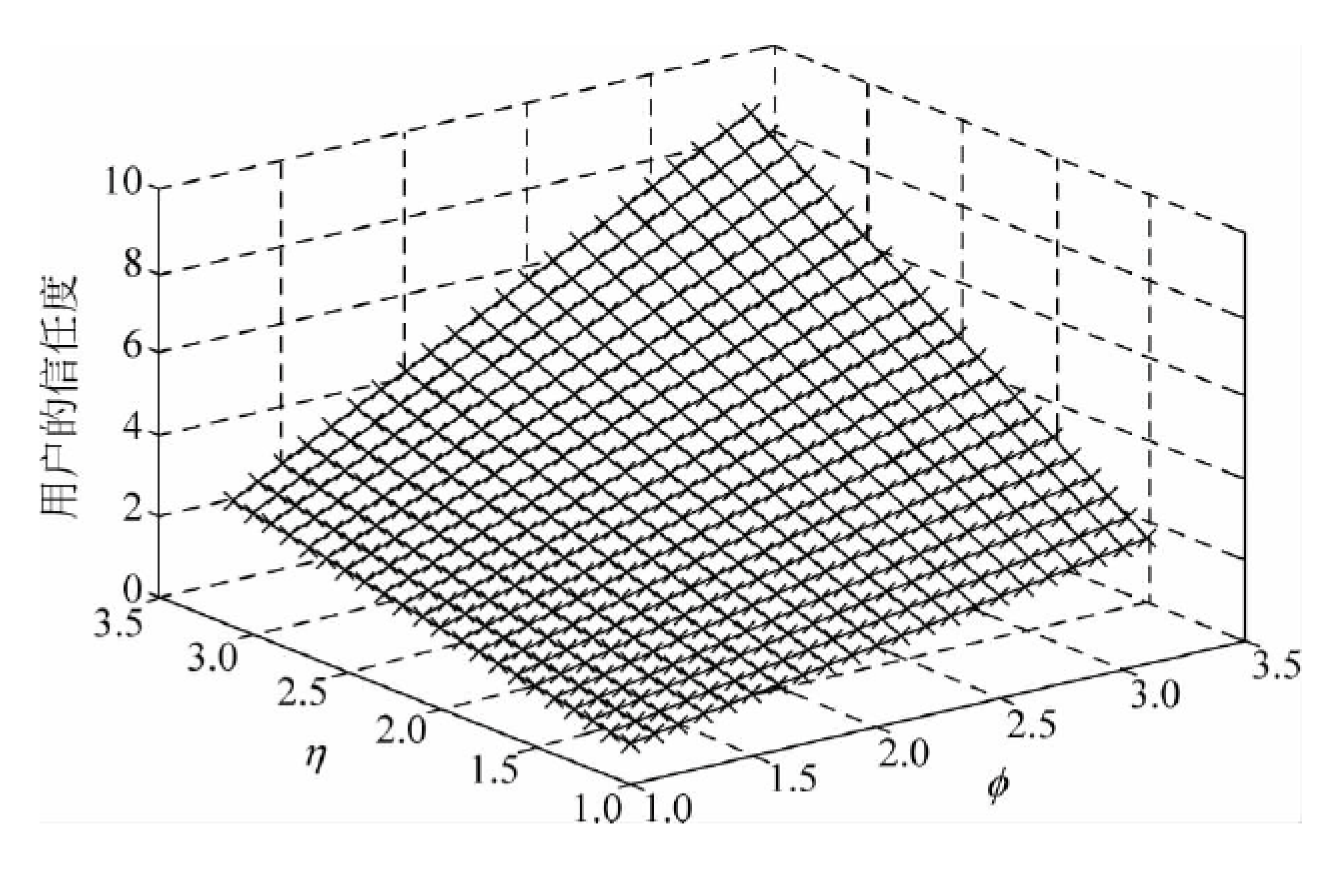


图 8-8　累积信任度分配

（4）用户证书披露因子和操作偏好对其效用的影响。

α_k、β_k 的取值超过一定阈值时，用户的信任度虽然会增加，但其安全效用下降。此时，用户需开始新一轮的认证过程。在图 8-9 中，当用户的 α_k、β_k 增大时，用户的证书和操作偏好泄露机会增加，认证协调器保持效用为一个常量，认证代理的效用随之增加，但用户的效用随之不断减小。当 $\alpha_k,\beta_k>3$ 时，用户的效用趋近于 0，此时用户虽然可获得认证代理分配的信任度，但其效用变到最低，因此，$\alpha_k,\beta_k>3$ 时，用户将不能再出示证书；否则将使得非法窃听者完全获得其证书和操作偏好。为了完成证书认证和防止非法用户的窃听，认证协调器此时将终止当前的认证过程，开始新的证书认证策略链，直到用户认证通过为止。

（5）α_k 和 c_m 对多个用户信任度分配的影响。

由式(8-21)可知，认证代理分配给用户的最优策略为 D_i^*，它根据披露因子 α_k 和利益函数 $c(\lambda')$ 来决定其分配策略。从图 8-10 中可以看出，当用户的证书披露因子 α_k 增大且利益函数常量 c_m 减小时，认证代理分配给用户的平均信任度在增长。例如，当 $\alpha_3=0.2$、$c_3=0.3$ 时，优化的平均信任度较低；当 $\alpha_1=0.9$、$c_1=0.1$ 时，优化的平均信任度最高，这意味着认证代理分配给用户高的信任度，这是由于证书披露因子 α_k 最高，但利益函数常数 c_m 最小时，根据式(8-21)，得分配给用户的信任度 D_i^* 值增大。

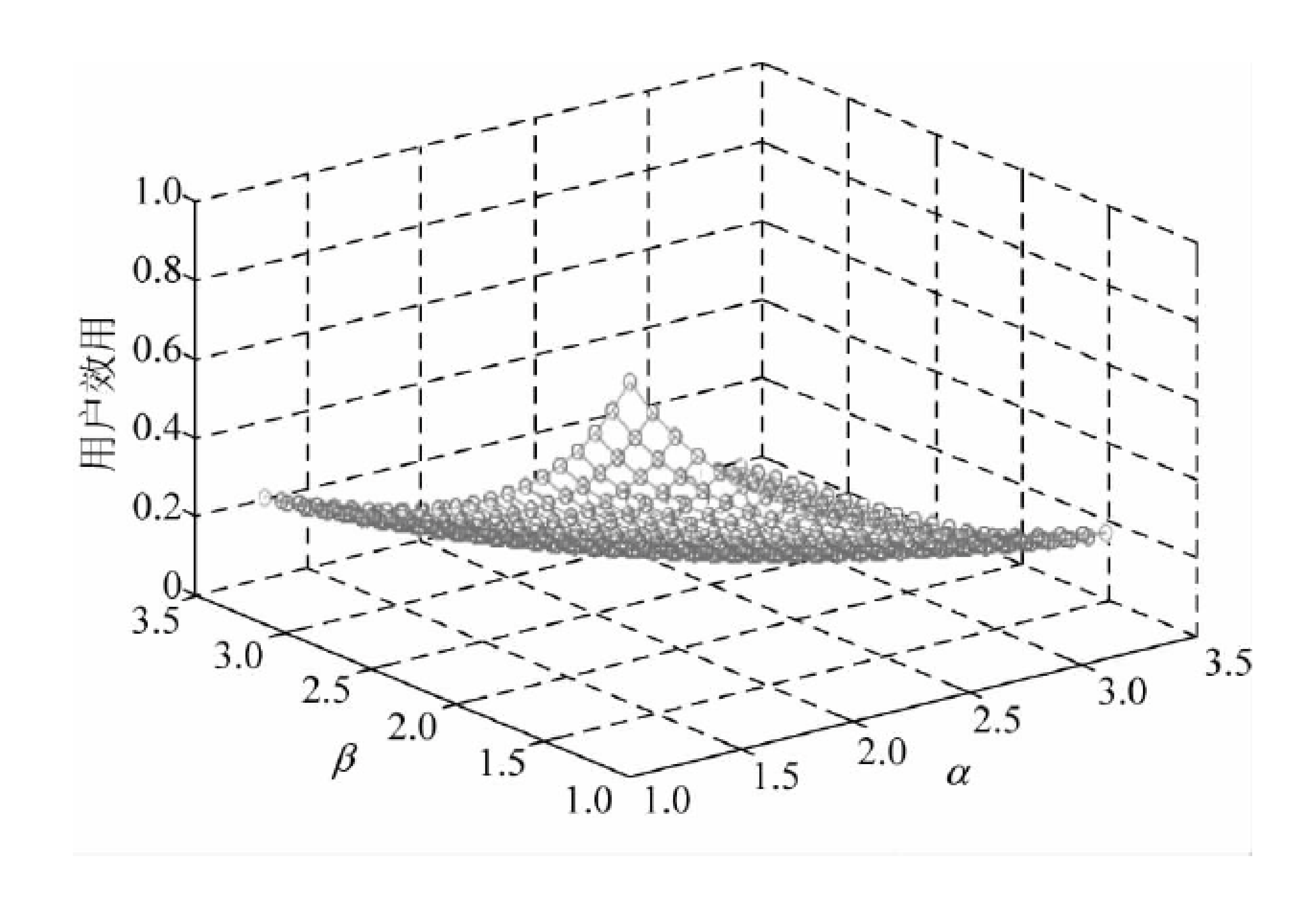


图 8-9 用户证书披露因子和操作偏好对其效用的影响

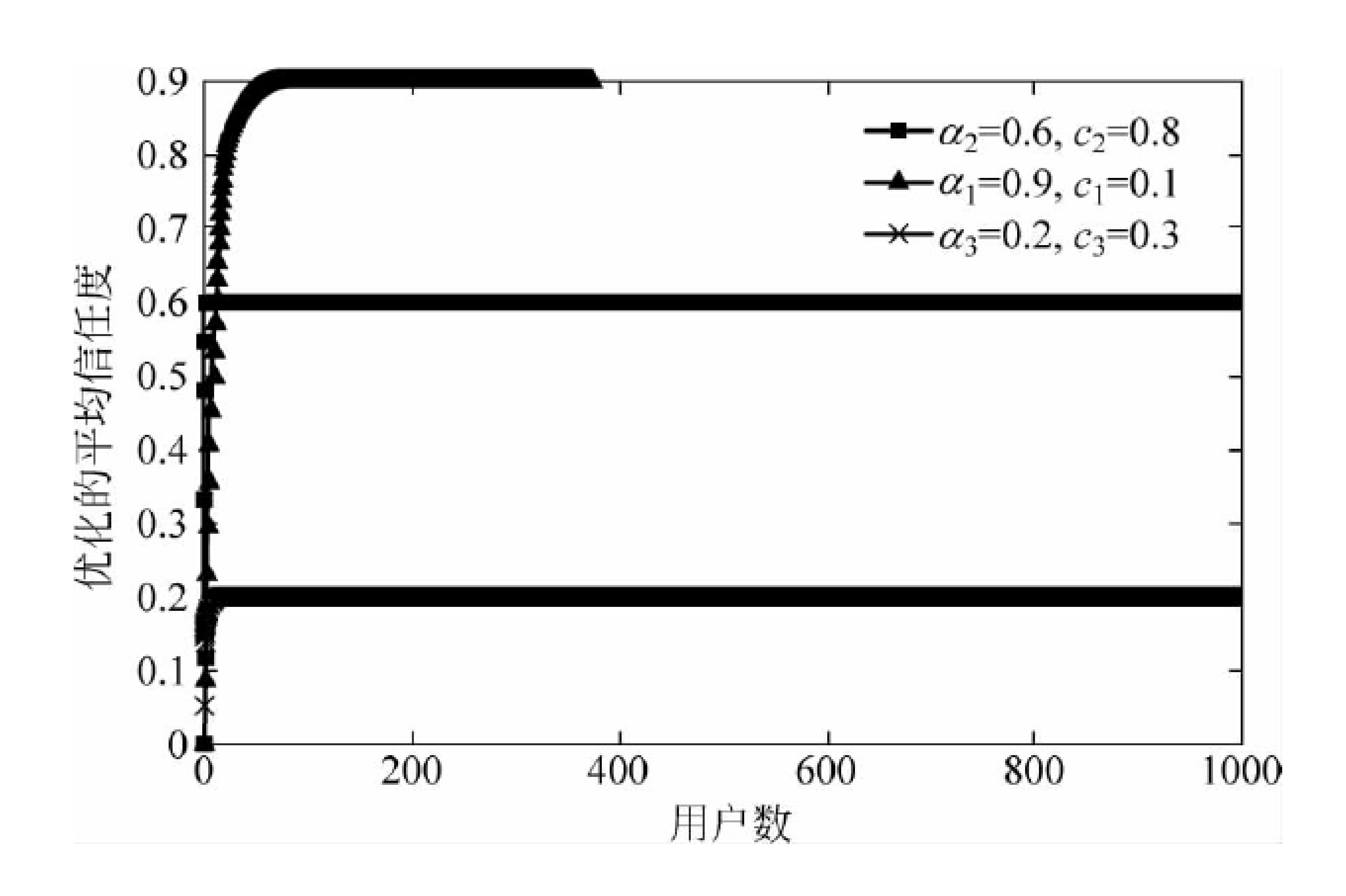


图 8-10 α_k 和 c_m 对多个用户信任度分配的影响

(6)学习增益对用户到认证代理、认证代理到认证协调器证书传递路径的信任协同学习误差的影响。

多个用户、认证代理、认证协调器经过多阶段博弈和多操作证书链验证才达到信任协同。达到信任协同稳定也即认证结束,信任协同时间影响了认证机制的性能,认证性能的好坏取决于学习增益对于用户到认证代理、认证代理到认证协调器证书传递路径的学习误差的影响。

图 8-11 给出了用户、认证代理和认证协调器之间学习增益对于信任协同学习误差的影响。使用非零的学习增益 $\mu_1=0.5$ 表示用户 S_i 的信任协同学习增益值,$\mu_2=0.7$ 表示在证书传递路径 $S_i\rightarrow AA_j$ 上认证代理 AA_j 的信任协同学习增益值,$\mu_3=0.1$ 表示在证书传递路径 $AA_j\leftarrow AC_v$ 上认证代理 AA_j 的信任协同学习增益值,$\mu_4=0.3$ 表示认证协调器的信任协同学习增益值。由于 $\mu_{ij,h}(t)=\mu_1+\mu_2$、$\mu_{vj,h}(t)=\mu_3+\mu_4$,所以 $\mu_{vj,h}(t)<\mu_{ij,h}(t)$ 表明认证代理和认证协调器之间的信任演化协同快于用户和认证代理之间的演化协同,越小的学习增益表明在很短的时间内信任达到稳定状态,可以极大提高认证系统的性能。

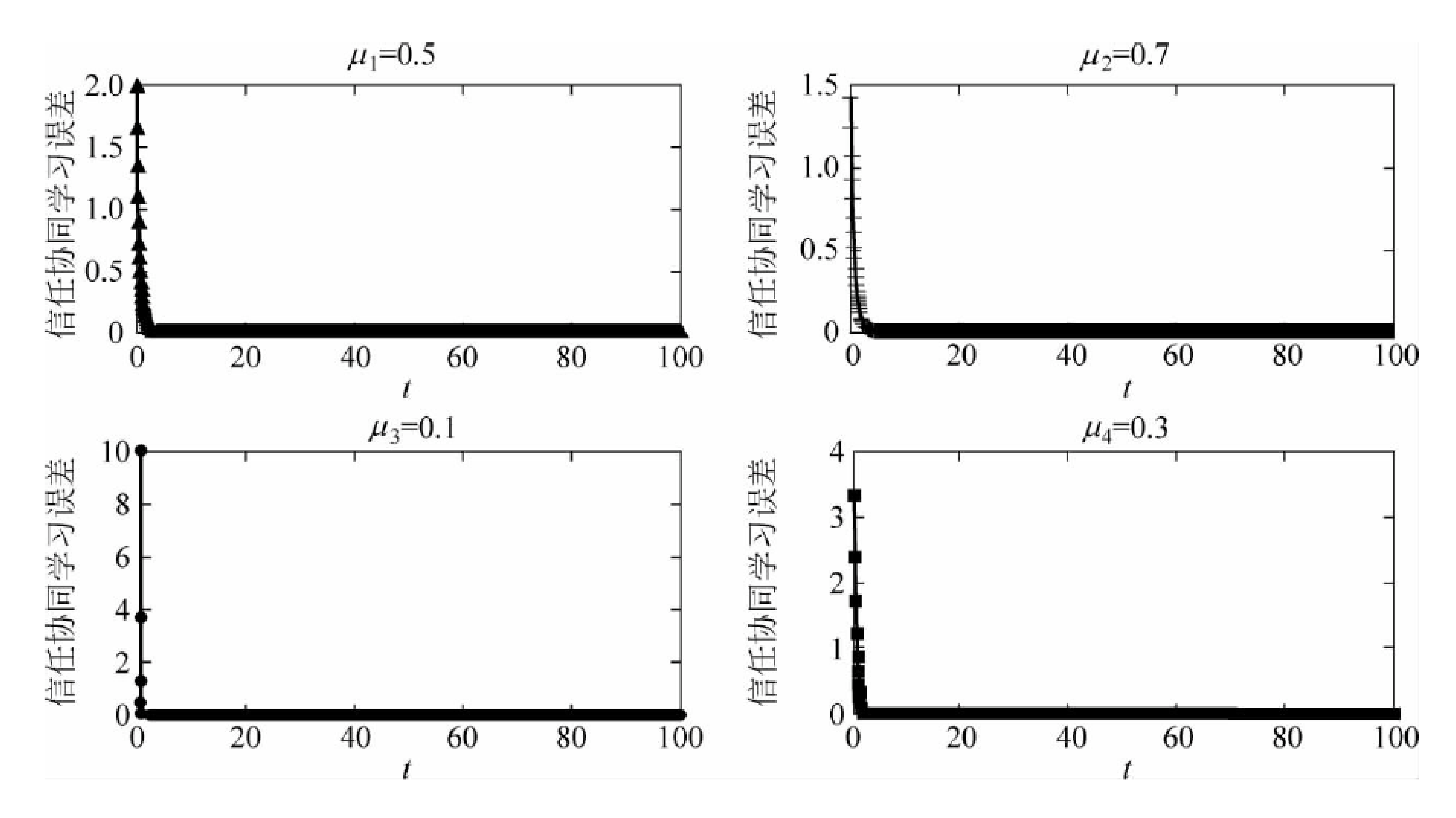


图 8-11 学习增益对信任协同学习误差的影响

（7）与传统的基于属性和本体角色的访问控制效用比较。

本章提出的证书博弈的动态认证机制考虑了证书泄露因子和操作偏好、动态证书策略链等因素，基于增益学习的信任协同有效地缩短了认证时间，提高了用户、认证代理和认证协调器之间的安全效用。

在图 8-12 中，与传统的基于属性和本体的访问控制系统进行了比较，当平均证书泄露度变化时，访问控制参数和学习增益 μ 使得证书的披露数自适应变化。当用户数增长时，动态证书认证演化博弈通过调节证书披露因子、操作偏好、学习增益使得认证协调器和用户的效用和信任度不断增强。在本实验中，当 $\alpha_k=0.9$，$c_m=0.1$，$\alpha_k+\beta_k<3$，$\mu_{vj,h}(t)<\mu_{ij,h}(t)$ 时，通过控制认证系统的性能参数，可使系统动态认证的信任和安全效用最优。对于传统的基于属性和本体的访问控制系统而言无法动态处理大量的传感云用户的认证请求，使得系统的效用下降，同时，信任度也很难建立。此外，由于在传感云计算环境中，传感云用户和资源之间的关系是 Ad Hoc 和动态的，资源和用户在不同的安全域。用户的身份经常由他们的特征和属性来识别，而不是使用证书预先定义用户身份。例如，对分布式多服务器体系结构常使用智能卡识别用户身份[415]。但是智能卡容易被伪造，而动态的证书不容易被伪造，

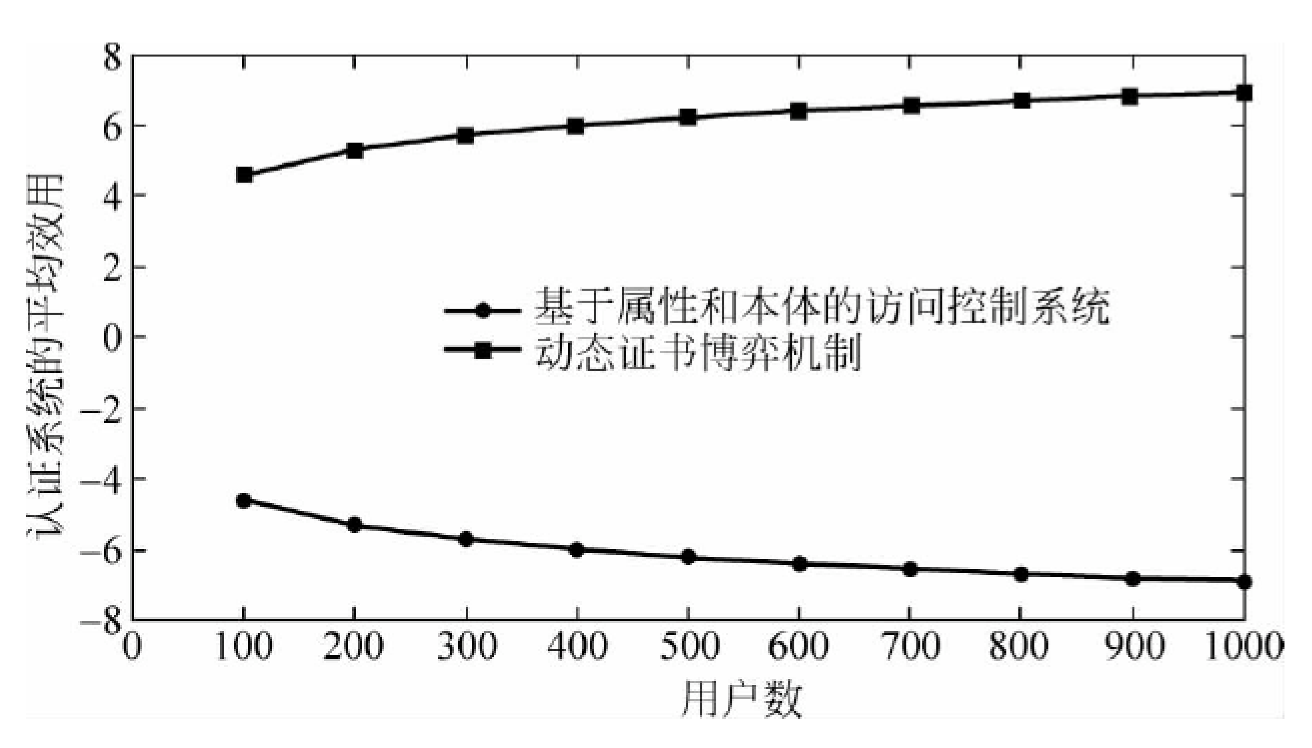


图 8-12 与传统的基于属性和本体的访问控制效用比较

即使被伪造了，还要经过多轮的认证博弈，所以能有效过滤恶意的攻击者。因此，传统的根据属性和本体来做访问控制决策[416]的模型对于传感云计算系统是不高效、不安全的。与传统的基于属性和本体[417]的静态访问控制机制相比，本章中提出的动态证书博弈机制提高了安全效用和认证性能。

8.8 小结

当用户通过披露证书执行传感云数据外包中心访问操作时，证书和操作偏好可能会泄露给窃听者。本章针对传感云数据外包中心访问控制系统提出了基于动态证书博弈的框架。证书认证博弈交互过程中，经过认证代理补偿一定的信任度来激励用户出示更多的证书，以提高其信任度。用户和认证协调器通过平衡证书泄露和信任补偿之间的关系来决定是否执行数据访问操作，认证代理根据用户披露的证书决定信任度的分配。本章模型化了证书认证的信任演化博弈的系统框架为一个3阶段的多轮博弈，使用多轮迭代博弈效用分析法分析了证书认证信任演化博弈的稳定性。通过数值仿真验证了用户、认证代理、认证协调器的优化策略，数值仿真显示，在传感云计算环境下，本章提出的证书认证信任演化博弈在效用和性能方面高于传统的基于属性和本体的访问控制系统。

第 9 章

基于随机演化联盟博弈的虚拟传感云服务安全机制研究

本章提出了随机演化联盟博弈框架并以此分析受攻击传感云服务系统的安全机制。在博弈的每一阶段，传感云服务提供者能够观察到服务组合节点的虚拟容量和攻击者的策略，根据这些观察，决定需分配的虚拟容量值来保证可靠安全的服务组合。传感云服务提供者通过 minimax-Q 和演化联盟形成学习算法，自适应地变化防御策略，对攻击者进行动态防御并形成可靠安全的服务组合。与随机博弈和演化联盟博弈相比，本章提出的随机演化联盟博弈策略在动态虚拟的安全服务组合过程中获得了较好的性能。

9.1 引言

传感云作为一个异构的网络通信环境，它利用云体系结构管理物理传感器节点，把物理传感器节点映射成虚拟传感节点，通过虚拟传感云服务处理物理传感器节点的感知数据。虚拟传感节点共享云资源，这使得各种物理传感器节点的数据能同时通过云平台来处理，极大地增强了物理传感器数据的处理速度。因此，传感云在环境监控、智慧农业、健康护理等方面具有诸多应用。

传感云平台作为云服务提供者管理和维护云服务，但是，要使得传感云服务系统安全可靠地运行还面临着诸多安全挑战。在传感云环境中，攻击者能够利用云服务系统的漏洞和资源来实施攻击。当前，入侵检测技术是一项用于解决虚拟传感云服务系统安全的有效技术。通过入侵检测技术能够监控攻击者的攻击策略以及系统的虚拟容量状况，据此，虚拟传感云服务提供者能够感知可以获得虚拟容量的节点，选择最佳的协作节点组合成可靠安全的服务网络，以此为私有云和公有云用户提供服务。

在云计算体系的多层结构中，基础设施、平台和应用都会遭到不同的安全威胁，恶意的攻击者可能在云计算系统的不同层上发起攻击。然而，已有的工作大都假设攻击者采用固定且不随时间变化的策略，如果攻击者也采用系统资源监控技术，它们很可能根据动态的网络环境和防御策略自适应地调整攻击策略。

本章主要讨论对传感云服务系统的容量攻击，并且提出随机演化联盟博弈的框架来设计虚拟传感云服务组合的可靠安全的防御策略，它能随虚拟传感云服务组合环境、服务质量、资源状态等系统的性能参数而变化。通过模型化策略和动态博弈把防御者和攻击者之

间的交互看作零和随机演化联盟博弈。为了确保虚拟传感云服务组合的可靠性、安全性和组合容量的最大化，提出了为私有云和公有云提供服务组合容量的混合分配机制，并且根据攻击者变化的策略自适应地分配容量。其中，定义的演化联盟博弈用于合作容量分配，而随机博弈用于观察攻击者的行动状态。

本章使用 minimax-Q 学习算法，为虚拟传感云服务提供者获得了最优策略。当云计算中心容量缺乏时，虚拟传感云服务提供者将减少公有云用户的容量，从而为私有云用户预留一定的容量。当云计算中心容量升高时，空闲的容量将分配给公有云用户。虚拟传感云服务提供者通过演化联盟组成服务组合时，若云服务提供者观察到由于受到攻击使得容量降低难以保证服务质量的情况，云服务提供者将采取混合策略以避免在下一时刻受到严重攻击时降低服务质量。在云计算中心，当有多个服务提供者时，通过演化联盟形成容量共享的服务组合，可获得较高的服务质量。这样，通过动态地形成私有云服务组合联盟和公有云服务组合联盟，调节内部和外部容量，将使攻击者难以决策。因此，虚拟传感云服务提供者使用随机演化联盟博弈能获得较高的防御收益。

在扩展作者前期工作[418]的基础上，本章的工作主要包括以下内容：

（1）为实现虚拟传感云服务组合的可靠性，通过建立形式化的随机演化联盟博弈模型来获得攻击防御的最优策略。优化模型考虑了攻击者的攻击策略和服务组合质量的动态变化，通过对云计算中心的资源监控和攻击者行动的观察自适应地调整其服务组合策略。

（2）把云服务提供者对服务组合状态的观察模型化为有限状态的马尔可夫链（Finite State Markov Chain，FSMC）。它能描述云服务提供者和攻击者之间的随机博弈状态，通过它可以计算出双方的收益。

（3）使用 minimax-Q 学习算法获得攻击防御的最优策略。通过 Shapley 值来使得合作的云服务提供者获得公平的收益分配，激励云服务提供者合作形成联盟。使用联盟形成学习算法实现云服务提供者的收益，再通过学习最优策略形成稳定的可靠联盟，从而增强虚拟传感云服务的可靠性、安全性和服务质量。

本章其余章节安排如下：9.2 节总结相关工作；9.3 节描述虚拟传感云服务安全框架；9.4 节描述传感云服务安全博弈模型；9.5 节阐述随机演化联盟博弈优化策略；9.6 节阐述随机演化联盟均衡学习策略；9.7 节是实验和仿真；9.8 节给出本章的小结。

本章涉及的符号含义如下：

G^s 表示虚拟传感云服务网络。

I 表示虚拟传感节点集合。

I_i 表示虚拟传感节点。

e_{ij} 表示虚拟传感节点 I_i 和 I_j 之间建立的虚拟链路，该链路在时刻 t 的服务质量为 $\mathrm{QoS}_{ij}(t)$。

$D_{ij}^{l}(t)$ 表示虚拟传感节点在正常通信情况下的通信时间。

$D_{ij}^{a}(t)$ 表示在受攻击情况下的通信时间。

$P_{ij}(t)$ 表示成功防御的概率。

$A_{ij}(t)$ 表示成功防御后获得的收益。

$C_{ij}(t)$ 表示虚拟链路 (I_i, I_j) 在时刻 t 可用的容量。

S' 表示由虚拟传感节点组成的服务组合的集合。

s_k'表示虚拟传感节点 I_i 可以提供功能为 f_i 的服务。

f_{ca}表示缓存数据功能。

f_{tr}表示转发数据功能。

$QoS_i^s(s_k')$表示虚拟传感节点 I_i 提供的服务 s_k'的服务质量。

$P_i^s(s_k')$表示使用 s_k'服务时防御成功的概率。

$A_i^s(s_k')$表示使用 s_k'服务时防御的收益。

$C_i^s(s_k')$表示节点 I_i 使用 s_k'服务时提供的容量。

QoS_{in}^s表示虚拟传感云内部服务提供者的服务质量。

QoS_{ex}^s表示虚拟传感云外部服务提供者的服务质量。

x_s 表示虚拟传感云内部服务提供者的数量。

y_s 表示虚拟传感云外部服务提供者的数量。

G_s 表示为每个传感云服务提供者分配的容量。

$C_{ij,m}$表示虚拟链路 e_{ij} 在联盟 $m \in M$ 中可用的容量。

$C_{ij,m}^d$表示虚拟链路 e_{ij} 在联盟 $m \in M$ 中正常传输时的容量。

$C_{ij,k}^a$表示虚拟链路 e_{ij} 在受攻击者 $k \in K$ 攻击时的容量。

r_t 表示在时刻 t 的收益。

γ^t表示在时刻 t 的贴现因子。

S^t 表示虚拟传感云服务在时刻 t 的状态集合。

N'表示博弈参与者集合。

A_m^i 表示博弈参与者 i 在联盟 $m \in M$ 中可用的行动集合。

P 表示虚拟传感云服务将时刻 t 的状态集合 S^t 转换为时刻 $t+1$ 的状态集合 S^{t+1}的转换概率。

U 表示攻击者和防御者获得的收益。

e_{ij}'表示参与者 i 和 j 之间具有合作关系。

k_i 表示节点度。

R_i 表示参与者 i 能提供的资源。

η_i 表示联盟中虚拟传感节点 i 的可靠性配置。

C_i 表示参与者 i 能为其邻居节点提供的容量。

P_i 表示参与者 i 选择参与者 j 作为合作者形成联盟的选择概率。

C_j 表示虚拟传感节点 i 的邻居节点 j 提供的容量。

τ 表示描述虚拟机硬件和软件的环境变量。

$\langle c \rangle$表示稳定联盟的平均容量。

α 表示参与者的容量因子。

$D_{m,i}$表示在联盟 m 中参与者 i 的收益。

I_i^c 表示参与者 i 能提供的共享容量。

I_s 表示参与者 i 的邻居节点提供的共享容量。

C_i^α 表示参与者 i 在容量因子 α 的影响下提供的共享容量。

C_s^α 表示参与者 i 的邻居节点在容量因子 α 的影响下提供的共享容量。

$\eta_s^c = 1$ 表示参与者 i 与邻居节点合作。

Z_i 表示参与者 i 当前的负载。

Z_i^t 表示在时刻 t 参与者 i 的负载能力。

β 表示负载因子。

temp 表示虚拟传感节点的临时集合。

ω_j 表示邻居节点 j 的负载阈值。

ϕ_j 表示邻居节点 j 的容量阈值。

λ_j 表示邻居节点 j 的收益阈值。

N 表示理性的虚拟传感节点组成的博弈参与者集合。

$\overline{N}$ 表示理性的虚拟传感节点(博弈参与者)数量。

Ψ 表示虚拟传感节点类型集合。

co 表示合作形成联盟。

de 表示不具有可靠的能力合作形成联盟。

S 表示随机演化联盟博弈状态空间的笛卡儿积。

S^d 表示随机博弈的状态空间。

$\overline{S^e}$表示演化联盟的状态空间。

A 表示随机演化联盟博弈参与者的行动状态空间的笛卡儿积。

a 表示随机博弈防御者的行动集合。

$\widetilde{N}$表示随机博弈参与者对于攻击者动态变化的攻击策略采取随机防御行动的个数。

$\overline{a}$表示演化联盟参与者的行动集合。

$\frown{N}$ 表示演化联盟参与者的行动个数。

δ 表示随机演化联盟博弈的收益。

δ_i 表示随机博弈参与者的收益。

R^d 表示随机博弈参与者的收益函数。

$\overline{\delta_i}$表示演化联盟博弈参与者的收益。

R^e 表示演化联盟博弈参与者的收益函数。

$\delta_t(M,\eta_M^i)$表示在时刻 t 为防御行动和演化联盟获得的平均收益。

η_j^i 表示虚拟传感节点 i 与可靠的虚拟传感节点 j 合作形成联盟。

$\mathbb{F}$表示虚拟传感节点 i 的期望收益。

$\eta_{M\setminus\{i\}}^i$表示在联盟 M 中节点 i 的邻居节点的可靠度。

$p_i^q(\eta_{M\setminus\{i\}}^i)$表示在联盟 M 中节点 i 与其他联盟成员合作的可靠性概率。

P_j^i 表示虚拟传感节点 i 观测到邻居节点 j 为可靠节点的概率。

η_j^c 表示虚拟传感节点与邻居节点 j 合作。

η_j 表示邻居节点 j 的可靠度。

$C_{ij,m}^{d,q}(M,\eta_M^i)$表示虚拟传感节点 i 的期望容量效用函数。

$C_{ij,k}^{a,q}(M,\eta_M^i)$表示虚拟传感节点 i 受攻击时的平均容量损失函数。

$|M|$表示联盟 M 中虚拟传感节点的个数。

$c_{ij,k}^a(M)$表示联盟 M 中虚拟链路 e_{ij}受攻击者 k 攻击时的容量损失。

P 表示随机演化联盟博弈的传递概率。

p 表示随机博弈的传递概率。

$\Delta(S^d)$表示状态空间 S^d 的概率分布。

$\bar{p}$ 表示演化联盟博弈的传递概率。

$\Delta(\overline{S^e})$表示状态空间$\overline{S^e}$的概率分布。

$\triangleright$表示演化联盟的偏好。

η_n^t 表示虚拟传感节点的状态。

$\widetilde{S}^t$表示在时刻 t 随机演化联盟博弈虚拟传感节点状态。

$\eta_{i,m}^t$表示在时刻 t 联盟 m 中虚拟传感节点 i 的可靠性。

$C_{i,m}^t$表示在时刻 t 联盟 m 中虚拟传感节点 i 的容量。

$H_{i,m}^t$表示在时刻 t 受攻击联盟 m 中虚拟传感节点 i 的容量。

$P^{ur\to r}(t+1)$表示虚拟传感节点由不可靠状态转移到可靠状态的概率。

$P^{r\to ur}(t+1)$表示虚拟传感节点由可靠状态转移到不可靠状态的概率。

p_i^d 表示虚拟传感节点 i 处于不可靠状态的概率。

g 表示转移到可靠状态时的容量收益。

c_0 表示转移到不可靠状态时的容量。

p_i^g 表示虚拟传感节点 i 获得收益的概率。

S^t 表示在时刻 t 随机演化联盟博弈服务组合的状态。

S_m^t 表示与虚拟传感节点 i 的可靠性和容量相关的状态。

$H_{m,\text{in}}^t$和 $H_{m,\text{ex}}^t$分别表示在时刻 t 服务组合 m 受到攻击的内部和外部服务数。

a^t 表示在时刻 t 联盟博弈参与者的防御行动集合。

$a_{i,\text{in1}}^t$和 $a_{i,\text{ex1}}^t$表示选择未攻击的虚拟传感节点 i 加入联盟后，分别分配的内部和外部容量。

$a_{i,\text{in2}}^t$和 $a_{i,\text{ex2}}^t$表示选择以前受攻击的虚拟传感节点 i 加入联盟后，分别分配的内部和外部容量。

a_h^t 表示在时刻 t 攻击者的行动集合。

$a_{i,h1}^t$和 $a_{i,h2}^t$分别表示在时刻 t 攻击者分别选择以前未攻击和受攻击的虚拟传感节点 i 攻击其容量。

a_m^t 表示在时刻 t 联盟 m 选择的行动。

$a_{m,\text{in1}}^t$和 $a_{m,\text{ex1}}^t$表示联盟 m 分别选择未攻击的内部服务和外部服务作为服务提供者。

$a_{m,\text{in2}}^t$和 $a_{m,\text{ex2}}^t$表示联盟 m 分别选择以前受攻击的内部服务和外部服务作为服务提供者。

$a_{m,h}^t$表示在时刻 t 攻击者对联盟 m 采取的行动。

$a_{m,h1}^t$和 $a_{m,h2}^t$表示在时刻 t 攻击者分别对联盟 m 中以前未受攻击和受攻击的服务发起攻击行动。

$p(C_{\text{in1}}^t, C_{\text{ex1}}^t \mid H_{m,\text{in}}^t, H_{m,\text{ex}}^t, a_m^t, a_{m,h1}^t)$表示在时刻 t 未受攻击的虚拟传感云服务在提供内部服务或外部服务时受攻击的概率。

C_{in1}^t和 C_{ex1}^t表示在时刻 t 未受攻击的虚拟传感节点提供给内部服务和外部服务的容量数。

C_{in2}^t 和 C_{ex2}^t 表示在时刻 t 受攻击的虚拟传感节点在时刻 $t+1$ 转换为可靠状态能提供给内部服务和外部服务的容量数。

$c_{m,1}^t$ 表示未攻击的容量。

c_i 表示虚拟传感云服务的总容量。

$p(C_{in2}^t, C_{ex2}^t \mid H_{m,in}^t, H_{m,ex}^t, a_m^t, a_{m,h2}^t)$ 表示受攻击的虚拟传感节点在时刻 t 提供内部服务或外部服务时受攻击的概率。

$c_{m,2}^t$ 表示虚拟传感云服务组合联盟 m 被攻击的容量。

$p(S^{t+1,d} \mid S^{t,d}, a^t, a_h^t)$ 表示 $\overline{\overline{N}}$ 个虚拟传感节点组成的 $\overline{M}$ 个服务组合的随机博弈状态传递概率。

$\bar{p}(M^{t+1} \mid M^t)$ 表示联盟结构从状态 M^t 转换到状态 M^{t+1} 的概率。

$P(S^{t+1,d}, \overline{S^{t+1,e}} \mid S^{t,d}, \overline{S^{t,e}}, a^t, \overline{a^t}, a_h^t)$ 表示随机演化联盟博弈的状态传递概率。

$S^{t,d}$ 表示在时刻 t 随机博弈的状态空间。

$\overline{S^{t,e}}$ 表示在时刻 t 演化联盟的状态空间。

$\overline{a^t}$ 表示在时刻 t 演化联盟参与者采取的行动。

$C(S^t, a^t, \overline{a^t}, a_h^t)$ 表示在每个博弈阶段获得的收益。

$B(S^t, a^t, \overline{a^t}, a_h^t)$ 表示联盟参与者使用防御策略形成可靠服务组合的容量配置。

$p^a(S^t, a^t, \overline{a^t}, a_h^t)$ 表示所有的虚拟传感节点受到攻击的概率。

C_{max} 表示虚拟传感节点的最大容量。

$Q'(S^t, a^t, \overline{a^t}, a_h^t)$ 表示随机博弈的 Q 函数。

$V(S^{t+1,d}, \pi^*)$ 表示随机博弈状态更新的值函数。

$\eta_{t+1}(\overline{S^{t+1,e}})$ 表示演化联盟的可靠性更新函数值。

$Q(S^t, a^t, \overline{a^t}, a_h^t)$ 表示随机演化联盟博弈的 Q 函数。

$\eta_t(\overline{S^{t,e}})$ 表示在时刻 t 的演化联盟的可靠性更新函数值。

$V(S^{t+1,d}, \eta_{t+1}(\overline{S^{t+1,e}}), \pi^*)$ 表示随机演化联盟博弈状态更新的值函数。

$C_i(\cdot)$ 表示虚拟服务组合联盟获得的收益。

π_i 表示联盟参与者 i 的策略。

π_i^* 表示联盟参与者 i 的纳什均衡策略。

π_{-i}^* 表示除联盟参与者 i 外的所有联盟参与者的纳什均衡策略。

$\varphi_i(c)$ 表示每个联盟参与者 i 的平均收益分配。

$|M^{in}|$ 表示向内提供服务的虚拟传感节点总数。

$|M^{ex}|$ 表示向外提供服务的虚拟传感节点总数。

$\varphi_j^{in}(c)$ 表示虚拟传感节点 j 向内提供服务时，获得 Shapley 值的平均收益分配概率。

$\varphi_j^{ex}(c)$ 表示虚拟传感节点 j 向外提供服务时，获得 Shapley 值的平均收益分配概率。

$c_t^j(S^t, a^t, \varphi_j(c))$ 表示随机演化联盟博弈中虚拟传感节点 j 的 Shapley 值平均收益分配的容量。

$c(M)$ 表示联盟 M 的平均容量。

$c_{t[0]}^{in}$ 和 $c_{t[1]}^{in}$ 分别表示内部服务不合作和合作获得的收益分配。

$c_{t[0]}^{ex}$ 和 $c_{t[1]}^{ex}$ 分别表示外部服务不合作和合作获得的收益分配。

$c_{m,t}^{j,\mathrm{in}}$表示在时刻 t 联盟参与者 j 选择内部服务组合联盟 m 后观察到的收益。

$\tilde{c}_{m,t}^{j,\mathrm{in}}$表示在时刻 t 内部服务组合联盟 m 中博弈参与者 j 期望的收益。

$\sigma_{j,t}^{\mathrm{in}}$表示在时刻 t 博弈参与者 j 在内部服务组合联盟的学习速率。

$\tilde{c}_{m,t}^{j,\mathrm{ex}}$表示在时刻 t 外部服务组合联盟 m 中博弈参与者 j 期望的收益。

$c_{m,t}^{j,\mathrm{ex}}$表示在时刻 t 联盟参与者 j 选择外部服务组合联盟 m 后观察到的收益。

$\sigma_{j,t}^{\mathrm{ex}}$表示在时刻 t 博弈参与者 j 在外部服务组合联盟的学习速率。

$\theta_{M,t}^{j}$表示若在时刻 t 虚拟传感节点 j 被选中加入到联盟 M，则其值为 1；否则为 0。

$\eta_{M,t}^{j}$表示若在时刻 t 虚拟传感节点 j 处于可靠状态，则其值为 1；否则为 0。

ξ_t 表示在时刻 t 博弈参与者 j 的学习速率。

$\Delta_j(t,t+1)$表示联盟参与者 j 变化其策略的概率。

$\Omega^*(S^*,P^*,\pi^*,\delta^*,M^*)$表示随机演化联盟博弈稳定状态。

S^* 表示随机演化联盟稳定的状态。

P^* 表示稳定状态的传递概率。

π^* 表示稳定策略。

δ^* 表示稳定状态的收益。

M^* 表示稳定联盟结构。

C_{tr} 表示博弈中联盟的通信成本。

Υ_i 表示服务组合联盟 i 被请求消息的次数。

ζ_i 表示消息在服务组合联盟 i 中传递的跳数。

d_r^i 表示服务组合联盟 i 请求消息的数据量。

d_p^i 表示服务组合联盟 i 接收消息的数据量。

9.2 相关工作

在云平台中，多个租户共享相同的虚拟服务，一方面，攻击者可通过 Cross-VM 攻击进入共享虚拟资源池获得用户的 RSA 和 AES 密钥；另一方面，攻击者能把正常的虚拟机(VM)替换为恶意的 VM 为用户所使用[419]，这对第三方的云计算平台安全构成了威胁。Zhang 等人[420]使用 HomeAlone 工具探测共享资源池中的异常活动。Santos 等人[421]为 IaaS 服务提出了信任的云计算框架，提供了封闭的云服务执行环境。由于数据中心网络预设的带宽比实际需求小，所以，DoS 攻击者可利用数据中心带宽的不足通过攻击影响网络正常通信。为了解决该问题，Liu 等人[422]提出了 DoS 避免策略，在云内部设置带宽监控代理，一旦检测到带宽下降，监控代理将执行应用迁移，有效地使用服务迁移技术处理泛洪攻击。为了防御 DoS 攻击，还有基于统计、数据挖掘、机器学习的方法，这些方法对于动态的云环境不具有较高的适应性。Girma 等人[423]提出了基于熵和协方差矩阵的方法分析 DoS 攻击，解决了当 DoS 攻击呈指数增长时的攻击探测问题。在文献[424，425]中，为了阻止 DoS 攻击，提出了可信的云存储服务模型，在这个模型中，为了不使数据完整性受到破坏，设计了一个分布式数据存储完整性审核机制。Kim 等人在文献[426]中研究云平台中的攻击类型，基于 OpenStack 工具仿真和测试了 SQL 注入、SYN-Flooding 攻击等。Arshad 等人在文献[427]中对虚拟机的入侵程度提出了一个入侵破坏评估机制，通过评估结果来决定防御

策略。Zhou 等人在文献[428]中针对协调攻击问题提出了协同入侵检测机制。为了探测不正常的活动，入侵检测系统能实时地了解检测对象的状态，但同时降低了系统性能，Kwon 等人在文献[429]中提出了入侵检测的自相似性测量方法，通过减少入侵检测系统的探测频率来提高入侵检测系统的性能。在文献[430，431]中作者研究了恶意节点的信任度探测和用户密码管理机制。Xie 等人在文献[432]中针对用户浏览 Web 网页的行为提出了大规模的隐式半马尔可夫链探测模型，还有研究者使用 Zipf 律测量 Web 页之间的相关性[433]来探测用户浏览网页的行为。大量的安全防御技术并不能为云平台提供预先警告，Kholidy 等人[434]提出了有限状态马尔可夫预测模型，使用自适应风险评估方法预测多阶段的云攻击。Zhang 等人[435]构建了一个小型的混合云系统，通过设计入侵检测机制来探测冷启动和 USB 自启动攻击。Chen 等人[436]针对多虚拟机环境恶意端口扫描行为，通过抽取日志中的行为记录进行入侵行为分析，使用攻击模型解决了攻击行为识别问题。

面对云计算平台中各种复杂的攻击行为，很难设计出满足一切需求的防御机制，Fan 等人[437]提出了随机博弈模型描述云计算中的攻击防御行为，使用 Petri 网验证了随机博弈模型进行攻击防御响应的正确性。Bedi 等人[438]针对不同节点共享虚拟机服务队列攻击提出了基于队列攻击的防御机制，但不能根据攻击者的恶意行为变化防御策略。Varadarajan 等人[439]针对公共云中的虚拟机资源攻击提出了修改虚拟机负载的防御机制，但只是修改了虚拟机负载，并没有结合虚拟机负载均衡来考虑防御机制。Zhou 等人[440]针对云计算中的调度器攻击提出了阻止攻击者唤醒虚拟机的方法，减少了攻击者对虚拟资源的占用。然而，调度器作为虚拟网络中的协调节点，如果攻击者采取的是直接使调度器出现故障的攻击行动，那么，阻止攻击者并唤醒虚拟机的防御机制将变得无效。Zhang 等人[441]针对侧信道攻击[69]使用 Vickrey-Clarke-Groves 博弈进行虚拟迁移。综上所述，恶意的攻击者可能在云计算系统的不同层上发起攻击，在云计算体系的多层结构中，基础设施、平台和应用均将遭受不同程度的安全威胁。

与以上的相关工作相比，本章主要讨论对传感云服务系统的容量攻击防御问题，提出了随机演化联盟博弈的框架来设计虚拟传感云服务组合的可靠安全的防御策略，首先在传感云环境下为虚拟传感云服务网络定义了一个防御模型，采用联盟结构描述虚拟传感云服务网络动态的通信场景。为了缩短联盟形成的时间和增加虚拟传感云服务组合的可靠性，运用 Barabasi-Albert(BA)模型形成可靠的联盟。虚拟传感节点在联盟中以存储—转发的方式合作分发传感数据，为了破坏传感数据正常的分发，恶意的攻击者通常在联盟中发起资源攻击，导致大量的虚拟传感节点变得不可靠。针对此问题，本章随后基于 Q-learning 和马尔可夫链技术开发了一个防御框架，把攻击者和防御者之间的策略变化模型化为一个随机博弈，使用马尔可夫链技术分析了它们之间的交互博弈过程。本章提出的随机演化联盟博弈模型使得虚拟传感云服务网络能有效地防御恶意攻击者的容量攻击。

9.3 虚拟传感云服务安全防御框架

9.3.1 虚拟传感云服务攻击模型

传感云框架使用数据中心的资源分发服务，当端用户请求服务时，虚拟传感节点自动组

成服务组合提供给端用户(End User)。由于传感云部署在开放平台中,恶意用户的请求对传感云的安全构成了威胁。图 9-1 给出了虚拟传感云服务攻击模型。在模型中,物理传感器节点通过 Map 函数映射成为虚拟传感节点,这些虚拟传感节点在演化联盟形成机制的作用下进行服务组合。端用户通过移动设备接入传感云平台请求数据,而恶意用户(Malicious User)通过扫描虚拟传感云服务漏洞,攻击其容量,从而使得服务组合失败或可用性降低。

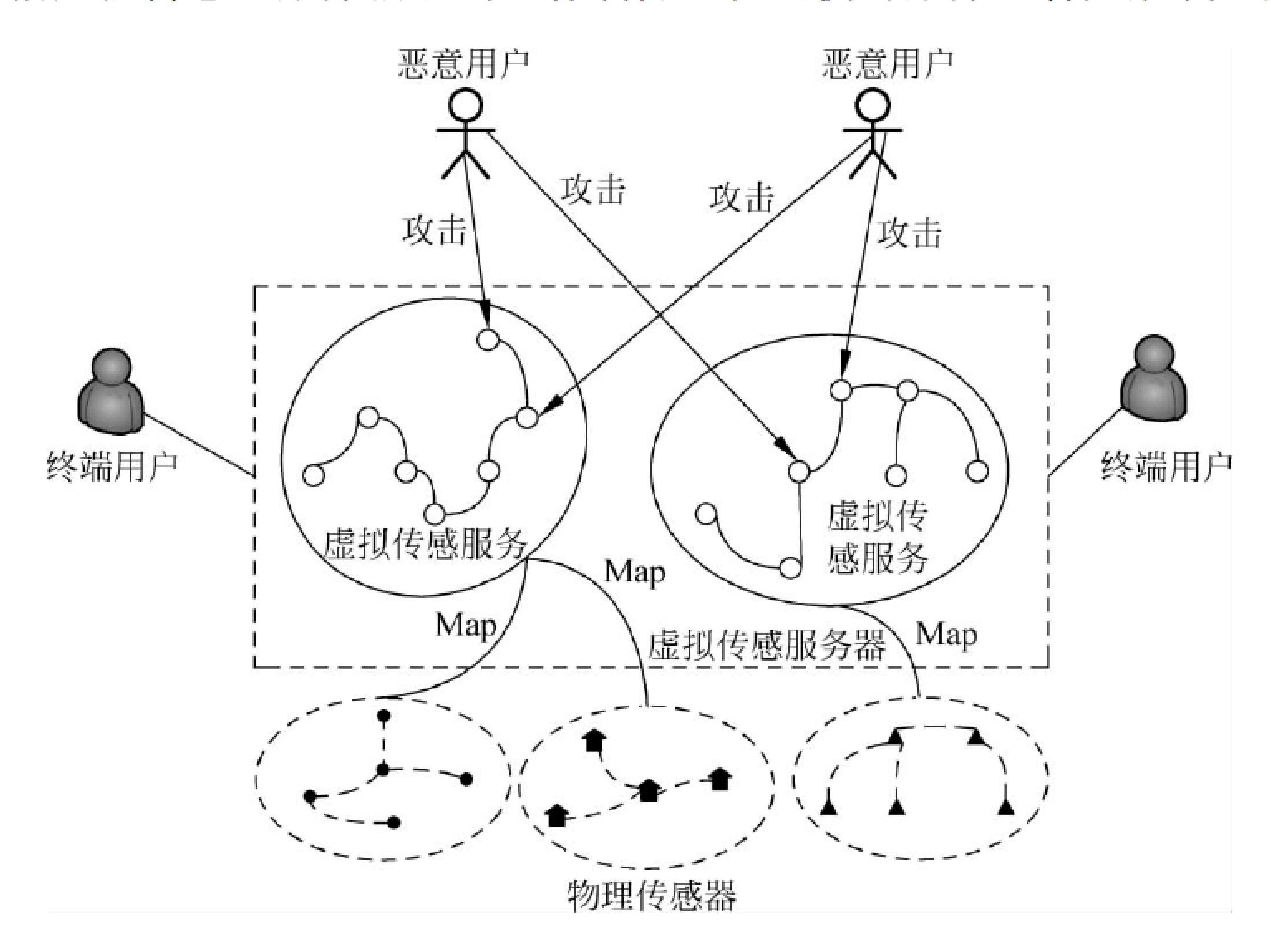


图 9-1 虚拟传感云服务攻击模型

定义 9-1 虚拟传感云服务网络由云用户和云服务提供者组成,用一个三元组 $G^s=(I,E,S)$ 表示,其中:

- $I=\{I_i \mid 1\leqslant i\leqslant |I|\}$ 表示虚拟传感节点集合,其中,虚拟传感节点定义为 $I_i=(\text{ints},\text{exts},\text{capacity})$,ints 表示向内提供服务,exts 表示向外提供服务。capacity 表示虚拟传感节点可用容量,如虚拟 CPU、虚拟内存、虚拟带宽等。
- $E=\{e_{ij} \mid e_{ij}\propto I_iI_j,1\leqslant i,j\leqslant |E|\}$ 表示虚拟传感云服务网络中的虚拟链路集合。其中,边 $e_{ij}=(I_i,I_j,\text{QoS}_{ij}(t))$ 表示虚拟传感节点 I_i 和 I_j 之间建立的虚拟链路,该链路在时刻 t 的服务质量为 $\text{QoS}_{ij}(t)$。$\text{QoS}_{ij}(t)=(D_{ij}^l(t),D(t)_{ij}^a,P_{ij}(t),A_{ij}(t),C_{ij}(t))$,其中,$D_{ij}^l(t)$ 表示虚拟传感节点在正常通信情况下的通信时间;$D_{ij}^a(t)$ 表示在受攻击情况下的通信时间;$P_{ij}(t)$ 表示成功防御的概率;$A_{ij}(t)$ 表示成功防御后获得的收益;$C_{ij}(t)$ 表示虚拟链路 (I_i,I_j) 在时刻 t 可用的容量。
- 服务组合 $S'=\{(I_i,s_k')\mid I_i\in I,I_i\rightarrow s_k'\}$ 表示由虚拟传感节点组成的服务组合的集合。其中,$s_k'=(I_i,f_i,\text{QoS}_i^s(s_k'))$ 表示虚拟传感节点 I_i 可以提供功能为 f_i 的服务。$f_i=\{f_{\text{ca}},f_{\text{tr}}\}$,其中,$f_{\text{ca}}$ 表示缓存数据功能,f_{tr} 表示转发数据功能。$\text{QoS}_i^s(s_k')=(P_i^s(s_k'),A_i^s(s_k'),C_i^s(s_k'))$ 表示虚拟传感节点 I_i 提供的服务 s_k' 的服务质量,其中,$P_i^s(s_k')$ 表示使用 s_k' 服务时防御成功的概率,且

$$P_i^s(s_k')=\frac{D_{ij}^l(t)}{D_{ij}^l(t)+D_{ij}^a(t)} \tag{9-1}$$

$A_i^s(s_k')$ 表示使用 s_k' 服务时防御的收益;$C_i^s(s_k')$ 表示节点 I_i 使用 s_k' 服务时提供的容量。

为了防御攻击者的攻击，保证可靠的虚拟服务组合，虚拟传感云服务提供者的最优策略是最大化服务质量。考虑私有云和公有云的混合传感云环境，除了虚拟传感云服务提供者要组合成内部服务为私有云用户提供服务外，还要组合成外部服务为公有云用户提供服务。因此，优化目标表示为

$$\max \sum_{s \in I} (\mathrm{QoS}_{\mathrm{in}}^{s} x_s + \mathrm{QoS}_{\mathrm{ex}}^{s} y_s)$$

$$\text{s. t. } \sum_{s \in I} (x_s + y_s) G_s \leqslant \sum_{i,j \in I} C_{ij,m}, x_s \geqslant 0, y_s \geqslant 0 \tag{9-2}$$

式中，$\mathrm{QoS}_{\mathrm{in}}^{s}$为虚拟传感云内部服务提供者的服务质量；$\mathrm{QoS}_{\mathrm{ex}}^{s}$为虚拟传感云外部服务提供者的服务质量；$x_s$ 为虚拟传感云内部服务提供者的数量；y_s 为虚拟传感云外部服务提供者的数量；G_s 为每个传感云服务提供者分配的容量；$C_{ij,m}$为虚拟链路 e_{ij} 在联盟 $m \in M$ 中可用的容量，可表示为

$$C_{ij,m} = (C_{ij,m}^{d} - \max C_{ij,k}^{a})^{+} \tag{9-3}$$

式中，$C_{ij,m}^{d}$为虚拟链路 e_{ij} 在联盟 $m \in M$ 中正常传输时的容量；$C_{ij,k}^{a}$为虚拟链路 e_{ij} 在受攻击者 $k \in K$ 攻击时的容量。

9.3.2 虚拟传感云服务安全防御框架

攻击者和防御者对虚拟传感云服务的攻击防御决策问题可以模型化为一个随机博弈过程，攻击者和防御者之间的交互过程可以模型化为一个马尔可夫决策处理过程。攻击者和防御者之间以离散的时刻进行交互，在每一个时刻，演化联盟中的虚拟传感云服务提供者与其他联盟成员合作形成联合防御行动，其典型的目标是最大化其收益 $\sum_{t=0}^{\infty} \gamma^t r_t$，其中，$r_t$ 表示在时刻 t 的收益，$\gamma^t (0 \leqslant \gamma^t \leqslant 1)$ 表示在时刻 t 的贴现因子，用于平衡短期收益与长期收益。

定义 9-2 虚拟传感云服务系统自适应防御框架为一个五元组(S^t, N', A_m^i, P, U)，其中：

- S^t 表示虚拟传感云服务在时刻 t 的状态集合。
- N'表示博弈参与者集合。
- A_m^i 表示博弈参与者 i 在联盟 $m \in M$ 中可用的行动集合。
- P 表示虚拟传感云服务将时刻 t 的状态集合 S^t 转换为时刻 $t+1$ 的状态集合 S^{t+1}的转换概率。
- U 表示攻击者和防御者获得的收益。

图 9-2 描述了虚拟传感云服务提供者自学习的自适应防御框架。当攻击者采用行动 a_h^t 攻击虚拟传感云服务时，虚拟传感云服务系统以概率 $p(S^{t+1} | S^t, a_h^t)$从状态 S^t 转换为 S^{t+1}状态，攻击者从而获得了收益 $u(s^t, a_h^t)$。虚拟传感云服务系统上的监控代理监测到这种状态变化后通知虚拟传感云服务提供者。多个虚拟传感云服务提供者通过自学习后形成联盟获得收益 $u^*(s^t, a_v^t)$，再求得最优策略 $\pi(\theta, a_v^t)$，并根据该最优策略以相应的概率值采取行动 a_v^{*t} 进行防御，使得虚拟传感云服务系统以概率 $p(S^{t+1} | S^t, a_v^t)$从状态 S^t 转换为S^{t+1}状态，虚拟传感云服务系统从而获得收益 $u(s^t, a_v^t)$。

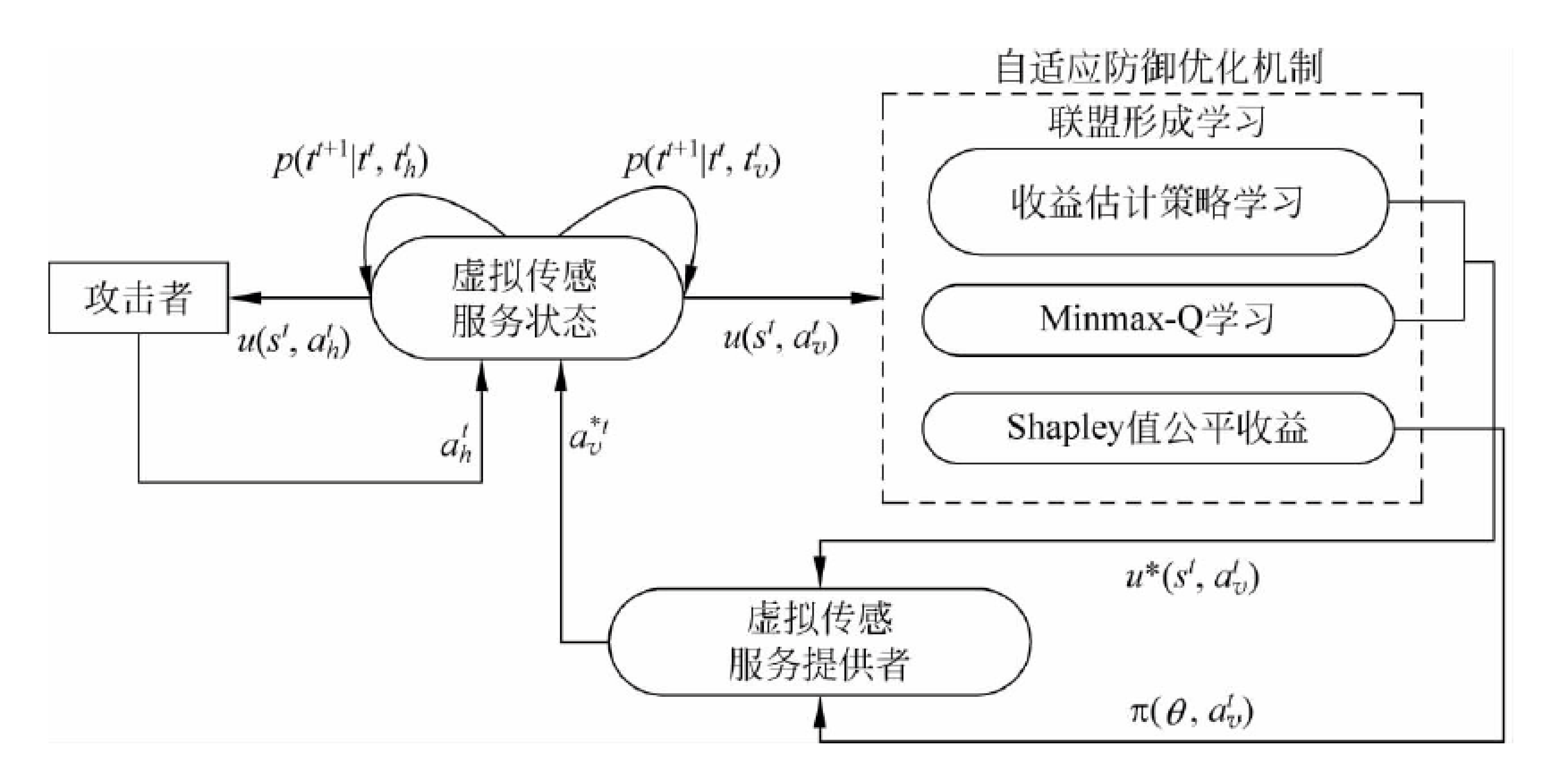


图 9-2　基于自学习的自适应防御框架

9.3.3　基于 BA 的随机演化联盟博弈模型

一个社会合作网络通常用节点和连接度信息进行描述，在本章建立的演化联盟博弈中，将每个虚拟传感节点看作一个社会网络节点，将虚拟传感节点间的联系看作共享容量度。因此，本章将利用 BA 无标度网络模型描述整个虚拟传感云服务系统的网络结构。设该传感云服务系统中有 N''个博弈参与者，用图 $G(V,E')$的顶点 $v_i \in V$ 表示参与者 i，无向边$e'_{ij} \in E'$表示参与者 i 和 j 之间具有合作关系。$C_i = k_i R_i$ 表示参与者 i 能为其邻居节点提供的容量，其中 k_i 为节点度，R_i 为参与者 i 能提供的资源。在虚拟传感云服务系统的网络模型中，每个参与者都有两种可配置的策略，$\Psi=\{\text{co},\text{de}\}$，其中，co 表示合作形成联盟，de 表示没有能力合作形成联盟。参与者 i 总是倾向于和可靠的虚拟传感节点合作形成联盟。

定义 9-3　联盟中虚拟传感节点 i 的可靠性配置定义为 $\eta_i = (C_i, P_i, D_{m,i}, Z_i)$，其中：

- C_i 表示参与者 i 能为其邻居节点提供的容量。
- $P_i = C_i / \sum_j C_j$ 表示参与者i 选择参与者j 作为合作者形成联盟的选择概率，其中，C_j 表示虚拟传感节点 i 的邻居节点 j 提供的容量。在联盟演化过程中，每个参与者根据自己的策略更新规则更新它的可靠性配置，经过迭代形成可靠的联盟。当参与者 i 更新它的策略时，它首先将随机地选择一个邻居节点 j，如果 $P_j > P_i$，参与者 i 将以概率 $p_{i\to j}$ 复制 j 的策略，其中

$$p_{i\to j} = \frac{1}{1 - \exp[(P_i - P_j)/\tau]} \tag{9-4}$$

式中，τ 为描述虚拟机硬件和软件的环境变量，它影响策略的复制和节点间的合作率。联盟参与者之间的合作率反映了联盟向稳定状态收敛的时间，合作率越大意味着联盟状态收敛于稳定的时间越短。联盟的合作率定义为

$$\upsilon = \left(1 - \frac{1}{\gamma^t}\right)(\langle c \rangle)^{\alpha} \tag{9-5}$$

式中，$\langle c \rangle$为稳定联盟的平均容量；α 为参与者的容量因子，$\alpha > 0$ 表示参与者 i 具有较高的容量，能提供较好的服务质量，而 $\alpha \leqslant 0$ 表示参与者 i 具有较低的容量，提供较差的服务质量。

- $D_{m,i}$表示在联盟 m 中参与者 i 的收益，定义为

$$D_{m,i}=\begin{cases}\left(I_i^c+\sum_s I_s\right)\times\dfrac{C_i^\alpha}{C_i^\alpha+\sum_s C_s^\alpha}, & \eta_s=1\\ 0, & \text{其他}\end{cases} \tag{9-6}$$

式中，

$$I_i^c=\frac{C_i^\alpha}{\sum_j C_j^\alpha} \tag{9-7}$$

为参与者 i 能提供的共享容量；I_s 为参与者 i 的邻居节点提供的共享容量；C_i^α 为参与者 i 在容量因子 α 的影响下提供的共享容量；C_s^α 为参与者 i 的邻居节点在容量因子 α 的影响下提供的共享容量；$\eta_s^c=1$ 表示参与者 i 与邻居节点合作。

- Z_i 表示参与者 i 当前的负载。负载与容量之间的关系可表示为

$$Z_i^t=\frac{C_i}{\beta} \tag{9-8}$$

式中，Z_i^t 为在时刻 t 参与者 i 的负载能力；β 为负载因子。如果参与者 i 当前的负载超过了预先定义的阈值 ω_i，则参与者 i 将不能被选择跟邻居节点形成联盟，参与者 i 成为联盟成员的概率为

$$\begin{cases}p_i(Z_i^t>\omega_i)=\dfrac{Z_i^t}{C_i}\\ p_i(Z_i^t\leqslant\omega_i)=Z_i^t\end{cases} \tag{9-9}$$

算法 9-1 基于 BA 模型的演化联盟形成算法。

1. 初始化联盟集合 $M_i=\varnothing$，temp$=\varnothing$；　//temp 表示虚拟传感节点的临时集合
2. FOREACH 虚拟传感节点 $i\in I$
3. 　temp=temp$\cup\{i\}$；
4. 　FOREACH 虚拟传感节点 $j\in I\backslash$temp
5. 　　IF $(Z_j^t<\omega_j)$ AND $(C_j>\phi_j)$ AND $(D_j>\lambda_j)$
　　　//ω_j 表示邻居节点 j 的负载阈值；ϕ_j 表示邻居节点 j 的容量阈值；
　　　//λ_j 表示邻居节点 j 的收益阈值。
6. 　　　虚拟传感节点 j 根据概率 P_i 加入虚拟传感节点 i 的联盟集合 M_i。
7. 　　　$M_i=M_i\cup\{(i,j)\}$。
8. 　　　$M_j=M_j\cup\{(j,i)\}$。
9. 　　ENDIF
10. 　　IF $(P_i>P_j)$
11. 　　　参与者 j 根据 $p_{i\to j}$ 复制虚拟传感节点 i 的策略。
12. 　　ENDIF
13. 　ENDFOR
14. ENDFOR
15. 使用联盟集合 M_i 创建图 $G(w,\varepsilon)$。
16. 设置图的顶点集 $w=I$。
17. 设置图的边集 $\varepsilon=\cup_{i=1}^{N''}M_i$。

通过基于 BA 模型的联盟形成算法可以过滤不可靠的虚拟传感节点，形成可靠的联盟。然而，在联盟工作阶段，由于已选择的邻居节点会受到攻击者的攻击，使得其工作状态变得不可靠，因此，参与者将重新选择虚拟传感节点组合成可靠的联盟。把由 BA 模型形成的可靠联盟作为一个虚拟的传感云服务为私有云和公有云用户提供服务，针对恶意用户发起的容量攻击建立随机演化联盟博弈模型进行防御。

9.4　虚拟传感云服务安全博弈模型

9.4.1　随机演化联盟博弈模型的防御策略分析

为防御虚拟传感云服务系统的攻击者，本章提出了一个随机演化联盟博弈模型(Stochastic Evolutionary Coalition Game，SECG)。在虚拟传感云服务系统中，攻击者实施服务攻击的目的是使虚拟资源耗尽，从而导致服务出现异常。而恶意用户对虚拟传感节点发起容量攻击的目的是阻止正常用户使用传感云服务系统的虚拟计算资源。这些受攻击的虚拟传感节点为私有云和公有云提供服务，因此，传感云服务系统的中心节点通过执行动态资源分配可以缓解潜在的攻击，防御者还可以通过动态的容量分配阻止攻击者。但是，使用容量分配来设计 SECG 模型中的防御策略时应考虑以下两个方面：

(1) 参与者在形成联盟时要平衡内部服务和外部服务的资源分配比例。传感云服务系统接收的外部服务的请求数一般大于内部服务的请求数。这样，外部服务的请求数越多，就需要分配给外部服务越多的容量，从而导致只能给内部服务分配较低的容量，此时，就需要平衡分配内部服务和外部服务之间的容量。

(2) 在面对攻击者变化的攻击策略时，参与者需要能自适应地调整防御策略。这是由于攻击者可能会在传感云服务系统上部署资源监控代理，根据监控代理获得的传感云服务系统资源状态和防御者的防御策略来动态调整其攻击策略。因此，防御者不能预先假设攻击者采取固定的攻击策略，而应该通过 SECG 模型来动态捕获攻击者的策略。

通过以上分析可以得出，防御者的目标是通过使用随机容量分配策略和形成演化联盟来提高虚拟传感节点的可靠性和虚拟传感节点服务组合的服务质量，而恶意攻击者的目标是通过攻击虚拟机的资源来降低虚拟传感节点的可靠性和服务组合的质量。它们具有相反的目标，因此，可以把它们之间动态的交互过程模型化为一个非合作的零和博弈，把虚拟传感云服务组合看作一个联盟，把虚拟传感节点看作防御者，而把恶意用户看作攻击者。此外，虚拟机的容量、虚拟传感云服务组合的服务质量、联盟结构及防御者和攻击者的策略是随时间变化的，所以，把虚拟传感云服务系统的安全防御机制形式化为一个随机演化联盟博弈模型。

9.4.2　随机演化联盟博弈模型的形式化定义

定义 9-4　随机演化联盟博弈模型结合了随机博弈和演化联盟博弈两种博弈类型。它可定义为一个七元组，即

$$\Omega = (N, \Psi, S, A, \delta, P, \triangleright) \tag{9-10}$$

其中：

- N 表示理性的虚拟传感节点组成的博弈参与者集合，其中包含 $\overline{N}$ 个理性的虚拟传

感节点(博弈参与者)。

- $\Psi=\{\text{co},\text{de}\}$表示虚拟传感节点类型集合,其中 co 表示合作形成联盟,de 表示不具有可靠的能力合作形成联盟。
- $S=S^d\times\overline{S^e}$表示随机演化联盟博弈状态空间的笛卡儿积,其中,$S^d$ 表示随机博弈的状态空间,$\overline{S^e}$表示演化联盟的状态空间。
- $A=a\times\overline{a}$表示随机演化联盟博弈参与者的行动状态空间的笛卡儿积,其中,$a=\{a_1,\cdots,a_{\widetilde{N}}\}$表示随机博弈防御者的行动集合,其中,$\widetilde{N}$表示随机博弈参与者对于攻击者动态变化的攻击策略采取随机防御行动的个数;$\overline{a}=\{\overline{a_1},\cdots,\overline{a_{\frown{N}}}\}$表示演化联盟参与者的行动集合,其中,$\frown{N}$ 表示演化联盟参与者的行动个数。
- $\delta=\delta_i\times\overline{\delta_i}$表示随机演化联盟博弈的收益。$\delta_i:S^d\times a_1\times\cdots\times a_{\widetilde{N}}\rightarrow R^d$ 表示随机博弈参与者的收益,R^d 表示随机博弈参与者的收益函数。$\overline{\delta_i}:\overline{S^e}\times\overline{a_1}\times\cdots\times\overline{a_{\frown{N}}}\rightarrow R^e$ 表示演化联盟博弈参与者的收益,R^e 表示演化联盟博弈参与者的收益函数。$\delta_t(M,\eta_M^i)$表示在时刻 t 防御行动和演化联盟获得的平均收益。$\eta_M^i=[\eta_1^1,\cdots,\eta_{\overline{\overline{m}}}^i]$,$\eta_j^i$ 表示虚拟传感节点 i 与可靠的虚拟传感节点 j 合作形成联盟。用整个联盟的平均收益表示单个节点期望的收益,结合式(9-3),虚拟传感节点 i 的期望收益为

$$\mathbb{F}=\sum_{t=0}^{\infty}\gamma^t\delta_t(M,\eta_M^i) \tag{9-11}$$

其中,

$$\begin{aligned}\delta_t(M,\eta_M^i)&=E[C_{ij,m}^d(M,\eta_M^i)-C_{ij,k}^a(M,\eta_M^i)]\\&=\sum_{q=1}^{2^{\overline{N}-1}}p_i^q(\eta_{M\setminus\{i\}}^i)(C_{ij,m}^{d,q}(M,\eta_M^i)-C_{ij,k}^{a,q}(M,\eta_M^i))\end{aligned} \tag{9-12}$$

式中,$\eta_{M\setminus\{i\}}^i$为在联盟 M 中节点 i 的邻居节点的可靠度;$p_i^q(\eta_{M\setminus\{i\}}^i)$为在联盟 M 中节点 i 与其他联盟成员合作的可靠性概率,定义为

$$p_i^q(\eta_{M\setminus\{i\}}^i)=\prod_{j\in M\setminus\{i\}}P_j^i(\eta_j^c=1\wedge\eta_j=\eta_j^i) \tag{9-13}$$

式中,P_j^i 为虚拟传感节点 i 观测到邻居节点 j 为可靠节点的概率;η_j^c 为虚拟传感节点与邻居节点 j 合作;η_j 为邻居节点 j 的可靠度。虚拟传感节点 i 的期望容量效用函数 $C_{ij,m}^{d,q}(M,\eta_M^i)$定义为

$$C_{ij,m}^{d,q}(M,\eta_M^i)=\begin{cases}Z_i(t)\beta, & e_{ij}>0\wedge x_{ij}=1\\0, & \text{其他}\end{cases} \tag{9-14}$$

式中,$e_{ij}>0$ 表示虚拟传感节点 i 与邻居节点 j 之间存在一条可靠边;$x_{ij}=1$ 表示虚拟传感节点 i 与邻居节点 j 通信;当两个虚拟传感节点通信时,所需动态容量被分配,当虚拟传感节点 i 受攻击时,平均容量损失函数 $C_{ij,k}^{a,q}(M,\eta_M^i)$定义为

$$C_{ij,k}^{a,q}(M,\eta_M^i)=\begin{cases}\sum_{i\in m,j\neq i}c_{ij,k}^a(M), & |M|>1\\0, & \text{其他}\end{cases} \tag{9-15}$$

式中,$|M|$为联盟 M 中虚拟传感节点的个数;$c_{ij,k}^a(M)$为联盟 M 中虚拟链路 e_{ij} 受攻击者 k 攻击时的容量损失。

- $P=p\times\overline{p}$表示随机演化联盟博弈的传递概率。$p:S^d\times a_1\times\cdots\times a_{\widetilde{N}}\rightarrow\Delta(S^d)$表示随

机博弈的传递概率，$\Delta(S^d)$表示状态空间 S^d 的概率分布；$\overline{p}:\overline{S^e}\times\overline{a_1}\times\cdots\times\overline{a_{\hat{N}}}\rightarrow\Delta(\overline{S^e})$表示演化联盟博弈的传递概率，$\Delta(\overline{S^e})$表示状态空间$\overline{S^e}$的概率分布。

- $\triangleright$表示演化联盟的偏好。例如，$M_1\triangleright_i M_2$ 表示参与者 i 倾向于加入联盟 M_2 并能获得较高收益和最优策略。

9.4.3　随机演化联盟博弈的状态和行动

假设在虚拟传感云服务系统中，$\hat{N}$个虚拟传感节点部署在虚拟服务器上。每个虚拟传感节点具有可靠和不可靠两种状态，用 η_n^t 表示虚拟传感节点的状态，其中，$\eta_n^t=1$ 表示虚拟传感节点在时刻 t 处于可靠的活跃状态，此时博弈参与者可以把它作为合作者组合成服务，而 $\eta_n^t=0$ 表示虚拟传感节点在时刻 t 处于不可靠状态，此时博弈参与者不能把它作为合作者组合成服务。虚拟传感节点的容量服务组合的 QoS 随时间变化，合作和防御策略也随时间变化。因此，随机演化联盟博弈可以模型化为一个有限状态的马尔可夫链（FSMC），这样虚拟传感节点的阶段防御的收益可以通过 FSMC 来描述。同时也注意到虚拟传感节点获得的容量收益依赖于自身的可靠状态。因此，为了增加容量收益，虚拟传感节点更倾向于选择可靠的节点合作形成联盟。在时刻 t 随机演化联盟博弈虚拟传感节点状态表示为$\widetilde{S}^t=\{\widetilde{S}_{1,m}^t,\cdots,\widetilde{S}_{\hat{b},m}^t\}$，其中，$\widetilde{S}_{i,m}^t=\{\eta_{i,m}^t,C_{i,m}^t,H_{i,m}^t\}$，$\eta_{i,m}^t$表示在时刻 t 联盟 m 中虚拟传感节点 i 的可靠性，$C_{i,m}^t$表示在时刻 t 联盟 m 中虚拟传感节点 i 的容量，$H_{i,m}^t$表示在时刻 t 受攻击联盟 m 中虚拟传感节点 i 的容量。

虚拟传感节点由不可靠状态转移到可靠状态的概率为

$$P^{ur\to r}(t+1)=p(\eta_{i,m}^{t+1}=1,C_{i,m}^{t+1}=g\mid\eta_{i,m}^t=0,C_{i,m}^t=c_0)=(1-p_i^d)p_i^g\quad(9\text{-}16)$$

虚拟传感节点由可靠状态转移到不可靠状态的概率为

$$P^{r\to ur}(t+1)=p(\eta_{i,m}^{t+1}=0,C_{i,m}^{t+1}=g-c_0\mid\eta_{i,m}^t=1,C_{i,m}^t=c_0)=p_i^d(1-p_i^g)\quad(9\text{-}17)$$

式中，p_i^d 为虚拟传感节点 i 处于不可靠状态的概率；g 为转移到可靠状态时的容量收益；c_0 为转移到不可靠状态时的容量，且 $c_0<\phi$，其中，ϕ 表示虚拟传感节点 i 的容量阈值；p_i^g 为虚拟传感节点 i 获得收益的概率。由式(9-16)和式(9-17)可以看出，虚拟传感节点可靠状态的变化影响了服务组合的容量和服务质量。如果一个联盟博弈参与者选择状态概率为 p_i^d 的虚拟传感节点来组合服务，那么，组合后的服务组合质量将降低，同时也不能保证虚拟传感云服务的安全性。因此，使用随机博弈和演化联盟博弈模型化虚拟传感云服务的安全防御为一个两阶段博弈来研究动态的攻击防御机制。在面对攻击者时，联盟博弈参与者选择虚拟传感节点进行合作，此时不仅考虑虚拟传感节点的可靠性，而且监测攻击者攻击策略的变化。对于攻击者而言，如果攻击者认为联盟参与者选择了一个可靠的虚拟传感节点形成联盟，则它将选择一个以前未被攻击的虚拟传感节点发起攻击行动。当一个虚拟传感节点通过防御策略分配了足够的容量时，它将以传递概率$(1-p_i^d)p_i^g$ 从不可靠状态切换到可靠状态，然后联盟博弈参与者选择这个可靠的节点作为联盟成员来形成服务组合。

定义 9-5　在时刻 t，随机演化联盟博弈服务组合的状态定义为 $S^t=\{S_1^t,\cdots,S_{\bar{c}}^t\}$，其中，$S_m^t=(\eta_{i,m}^t,C_{i,m}^t,H_{m,\text{in}}^t,H_{m,\text{ex}}^t)$表示与虚拟传感节点 i 的可靠性和容量相关的状态，其中，$H_{m,\text{in}}^t$ 和 $H_{m,\text{ex}}^t$ 分别表示在时刻 t 服务组合 m 受到攻击的内部和外部服务数。

定义 9-6 在时刻 t,联盟博弈参与者的防御行动集合定义为 $a^t=(a_1^t,a_2^t,\cdots,a_{\bar{d}}^t)$,其中,$a_i^t=(a_{i,\text{in1}}^t,a_{i,\text{ex1}}^t,a_{i,\text{in2}}^t,a_{i,\text{ex2}}^t)$,$a_{i,\text{in1}}^t$ 和 $a_{i,\text{ex1}}^t$ 表示选择未攻击的虚拟传感节点 i 加入联盟后分别分配的内部和外部容量。$a_{i,\text{in2}}^t$ 和 $a_{i,\text{ex2}}^t$ 表示选择以前受攻击的虚拟传感节点 i 加入联盟后分别分配的内部和外部容量。

定义 9-7 在时刻 t,攻击者的行动集合定义为 $a_h^t=\{a_{1,h}^t,a_{2,h}^t,\cdots,a_{\bar{e},h}^t\}$,其中 $a_{i,h}^t=(a_{i,h1}^t,a_{i,h2}^t)$,$a_{i,h1}^t$ 和 $a_{i,h2}^t$ 分别表示在时刻 t 攻击者分别选择以前未攻击和受攻击的虚拟传感节点 i 攻击其容量。

以上定义的行动选择集合给出了攻击者和联盟博弈参与者彼此间不确定的攻击和防御策略,其中,防御者的行动随着随机演化联盟博弈的状态发生变化,接下来将使用马尔可夫链技术来分析联盟参与者的博弈状态。

9.4.4 基于马尔可夫链的随机演化联盟博弈状态分析

在时刻 t 虚拟传感云服务组合联盟 m 面对攻击者攻击时,联盟 m 选择的行动为 $a_m^t=(a_{m,\text{in1}}^t,a_{m,\text{ex1}}^t,a_{m,\text{in2}}^t,a_{m,\text{ex2}}^t)$,其中,$a_{m,\text{in1}}^t$ 和 $a_{m,\text{ex1}}^t$ 表示联盟 m 分别选择未攻击的内部服务和外部服务作为服务提供者。$a_{m,\text{in2}}^t$ 和 $a_{m,\text{ex2}}^t$ 表示联盟 m 分别选择以前受攻击的内部服务和外部服务作为服务提供者。在时刻 t 攻击者对联盟 m 采取的行动为 $a_{m,h}^t=(a_{m,h1}^t,a_{m,h2}^t)$,其中 $a_{m,h1}^t$ 和 $a_{m,h2}^t$ 表示在时刻 t 攻击者分别对联盟 m 中以前未受攻击和受攻击的服务发起攻击行动。由于在下一个时刻 $t+1$,内部或外部服务将受到攻击,联盟参与者要从以前未攻击的或受攻击的虚拟传感节点集合中选出联盟成员组成服务组合。在时刻 t 未受攻击的虚拟传感云服务在提供内部服务或外部服务时受攻击的概率为

$$p(C_{\text{in1}}^t,C_{\text{ex1}}^t \mid H_{m,\text{in}}^t,H_{m,\text{ex}}^t,a_m^t,a_{m,h1}^t)=\frac{\begin{pmatrix}a_{m,\text{in1}}^t\\C_{\text{in1}}^t\end{pmatrix}\begin{pmatrix}a_{m,\text{ex1}}^t\\C_{\text{ex1}}^t\end{pmatrix}\begin{pmatrix}c_{m,1}^t-a_{m,\text{in1}}^t-a_{m,\text{ex1}}^t\\a_{m,h1}^t-C_{\text{in1}}^t-C_{\text{ex1}}^t\end{pmatrix}}{\begin{pmatrix}c_{m,1}^t\\a_{m,h1}^t\end{pmatrix}} \tag{9-18}$$

式中,C_{in1}^t、C_{ex1}^t 为在时刻 t 未受攻击的虚拟传感节点分别提供给内部服务和外部服务的容量数;C_{in2}^t、C_{ex2}^t 为在时刻 t 受攻击的虚拟传感节点在时刻 $t+1$ 转换为可靠状态能分别提供给内部服务和外部服务的容量数;$H_{m,\text{in}}^{t+1}=C_{\text{in1}}^t+C_{\text{in2}}^t$;$H_{m,\text{ex}}^{t+1}=C_{\text{ex1}}^t+C_{\text{ex2}}^t$;$c_{m,1}^t=c_i-H_{m,\text{in}}^t-H_{m,\text{ex}}^t$ 表示未攻击的容量,其中,c_i 为虚拟传感云服务的总容量。

类似地,受攻击的虚拟传感节点在时刻 t 提供内部服务或外部服务时受攻击的概率为

$$p(C_{\text{in2}}^t,C_{\text{ex2}}^t \mid H_{m,\text{in}}^t,H_{m,\text{ex}}^t,a_m^t,a_{m,h2}^t)=\frac{\begin{pmatrix}a_{m,\text{in2}}^t\\C_{\text{in2}}^t\end{pmatrix}\begin{pmatrix}a_{m,\text{ex2}}^t\\C_{\text{ex2}}^t\end{pmatrix}\begin{pmatrix}c_{m,2}^t-a_{m,\text{in2}}^t-a_{m,\text{ex2}}^t\\a_{m,h2}^t-C_{\text{in2}}^t-C_{\text{ex2}}^t\end{pmatrix}}{\begin{pmatrix}c_{m,2}^t\\a_{m,h2}^t\end{pmatrix}} \tag{9-19}$$

式中,$c_{m,2}^t=H_{m,\text{in}}^t+H_{m,\text{ex}}^t$ 表示虚拟传感云服务组合联盟 m 被攻击的容量。

在演化联盟中,每个参与者采取的防御行动是独立的,并且在攻击者的攻击下其状态是动态变化的,$\bar{\bar{N}}$ 个虚拟传感节点组成的 $\bar{M}$ 个服务组合的随机博弈状态传递概率表示为

$$p(S^{t+1,d} \mid S^{t,d},a^t,a_h^t)=\prod_{m=1}^{\bar{M}}\left\{p(H_{m,\text{in}}^{t+1},H_{m,\text{ex}}^{t+1} \mid H_{m,\text{in}}^t,H_{m,\text{ex}}^t,a_m^t,a_{m,h}^t)\times\sum_{i=1}^{\bar{\bar{N}}}p(\eta_{i,m}^{t+1},C_{i,m}^{t+1} \mid \eta_{i,m}^t,C_{i,m}^t)\right\} \tag{9-20}$$

式中，

$$p(H_{m,\text{in}}^{t+1},H_{m,\text{ex}}^{t+1}\mid H_{m,\text{in}}^{t},H_{m,\text{ex}}^{t},a_m^t,a_{m,h}^t)=p(C_{\text{in1}}^t,C_{\text{ex1}}^t\mid H_{m,\text{in}}^{t},H_{m,\text{ex}}^{t},a_m^t,a_{m,h1}^t)\times p(C_{\text{in2}}^t,C_{\text{ex2}}^t\mid H_{m,\text{in}}^{t},H_{m,\text{ex}}^{t},a_m^t,a_{m,h2}^t)\quad(9\text{-}21)$$

$p(\eta_{i,m}^{t+1},C_{i,m}^{t+1}\mid\eta_{i,m}^{t},C_{i,m}^{t})$可由式(9-16)和式(9-17)求得。联盟中的每个参与者在采取防御行动的同时，联盟参与者选择可靠的虚拟节点形成新的联盟使得联盟不断演化。联盟结构从状态M^t转换到状态M^{t+1}的概率表示为

$$\bar{p}(M^{t+1}\mid M^t)=\begin{cases}\displaystyle\sum_{i\in M^t}\frac{1}{N}\cdot\frac{1}{|M^t\backslash M_k^i|}\cdot\mathbf{1}_{\{M_j\triangleright_i M_k\}}, & M^{t+1}\neq M^t\\ \displaystyle 1-\sum_{i\in M^t}\frac{1}{N}\cdot\frac{1}{|M^t\backslash M_k^i|}\cdot\mathbf{1}_{\{M_j\triangleright_i M_k\}}, & M^{t+1}=M^t\end{cases}\quad(9\text{-}22)$$

式中，$\frac{1}{|M^t\backslash M_k^i|}$为一个联盟参与者$i$被选择加入联盟$M_k\in M^t\backslash M_k$的概率；$\mathbf{1}_{\{M_j\triangleright_i M_k\}}$为指示函数，指示参与者$i$被联盟$M_k$选中后，离开联盟$M_j$加入到联盟$M_k$的指示信息，如果$M_j\triangleright_i M_k$，则指示函数的值为1；否则为0。这使得联盟$M_k$的结构发生$M^t$到$M^{t+1}$的变化，结合式(9-20)和式(9-22)可得随机演化联盟博弈的状态传递概率为

$$P(S^{t+1,d},\overline{S^{t+1,e}}\mid S^{t,d},\overline{S^{t,e}},a^t,\overline{a^t},a_h^t)=p(S^{t+1,d}\mid S^{t,d},a^t,a_h^t)\times\bar{p}(M^{t+1}\mid M^t)\quad(9\text{-}23)$$

式中，$S^{t,d}$为在时刻t随机博弈的状态空间；$\overline{S^{t,e}}$为在时刻t演化联盟博弈的状态空间；$\overline{a^t}$为在时刻t演化联盟参与者采取的行动。

9.4.5　随机演化联盟博弈收益

随机演化联盟博弈的目标是最大化虚拟传感云服务的可靠性和服务质量。在每个博弈阶段获得的收益为

$$C(S^t,a^t,\overline{a^t},a_h^t)=B(S^t,a^t,\overline{a^t},a_h^t)(1-p^a(S^t,a^t,\overline{a^t},a_h^t))\quad(9\text{-}24)$$

式中，$B(S^t,a^t,\overline{a^t},a_h^t)$为联盟参与者使用防御策略形成可靠服务组合的容量配置，有

$$B(S^t,a^t,\overline{a^t},a_h^t)=\left\{1-\sum_{m=1}^{\bar{M}}\left[\frac{a_{m,h1}^t}{c_{m,1}^t}(a_{m,\text{in1}}^t+a_{m,\text{ex1}}^t)+\frac{a_{m,h2}^t}{c_{m,2}^t}(a_{m,\text{in2}}^t+a_{m,\text{ex2}}^t)\right]\right\}\sum_{m=1}^{\bar{M}}\sum_{i=1}^{\bar{N}}D_{m,i}\quad(9\text{-}25)$$

式中，$D_{m,i}$的值由式(9-6)求得；$p^a(S^t,a^t,\overline{a^t},a_h^t)$为所有的虚拟传感节点受到攻击的概率，它可表示为

$$p^a(S^t,a^t,\overline{a^t},a_h^t)=\frac{\binom{c_{m,1}^t-a_{m,\text{in1}}^t-a_{m,\text{ex1}}^t}{a_{m,h}^t-C_{\text{in1}}^t-C_{\text{ex1}}^t}}{\binom{c_{m,1}^t}{a_{m,h1}^t}}\times\frac{\binom{c_{m,2}^t-a_{m,\text{in2}}^t-a_{m,\text{ex2}}^t}{a_{m,h}^t-C_{\text{in2}}^t-C_{\text{ex2}}^t}}{\binom{c_{m,2}^t}{a_{m,h2}^t}}\quad(9\text{-}26)$$

9.5　随机演化联盟博弈优化策略

一般来说，任何一个虚拟传感节点有最大容量，期望的收益不能超过其最大容量。同时，获得的收益不能比平均收益小。联盟参与者使用随机演化联盟博弈的目标是求出最优

策略从而最大化期望收益，即

$$\max E\left\{\lim_{x\to\infty}\sum_{t=1}^{x}\gamma^t C(S^t,a^t,\overline{a^t},a_h^t)\right\}$$

$$\text{s.t. } \delta_t(M,\eta_M^i)\leqslant C(S^t,a^t,\overline{a^t},a_h^t)<C_{\max} \tag{9-27}$$

式中，$C_{\max}$为虚拟传感节点的最大容量。为了获得联盟参与者最优策略，使用 minimax-Q learning 算法，在时刻 t，当联盟参与者采取 $a^t=a_i^t\times\overline{a^t}$行动时，攻击者采用的行动为 a_h^t，由 Q 函数估计累积的收益。因此，联盟参与者通过更新 Q 函数值能计算纳什均衡策略并把 $Q(S^t,a^t,\overline{a^t},a_h^t)$的值作为博弈矩阵的收益。其中，随机博弈的 Q 函数为

$$\begin{aligned}&Q'(S^t,a^t,\overline{a^t},a_h^t)\\&=C(S^t,a^t,\overline{a^t},a_h^t)+\gamma^t\sum_{S^{t,d},S^{t+1,d}\in S^d,\overline{S^{t,e}},\overline{S^{t+1,e}}\in\overline{S^e}}p(S^{t+1,d},\overline{S^{t+1,e}}\mid S^{t,d},\overline{S^{t,e}},a^t,\overline{a^t},a_h^t)V(S^{t+1,d},\pi^*)\end{aligned} \tag{9-28}$$

式中，$V(S^{t+1,d},\pi^*)$为随机博弈状态更新的值函数，通过估计值函数获得随机博弈的最优策略 π^*。

在攻击者的攻击下，随机演化联盟博弈中的虚拟传感节点的状态变化决定了演化联盟的可靠性，演化联盟的可靠性更新函数值为

$$\begin{aligned}\eta_{t+1}(\overline{S^{t+1,e}})&=\frac{\sum_{\overline{S^{t,e}}\in\overline{S^e},i\in I}P^{w\to r}(t)\eta_{i,m}^t(\overline{S^{t,e}})}{\sum_{\overline{S^{t,e}}\in\overline{S^e},i\in I}P^{w\to r}(t)\eta_{i,m}^t(\overline{S^{t,e}})+\sum_{\overline{S^{t,e}}\in\overline{S^e},i\in I}P^{r\to w}(t)\eta_{i,m}^t(\overline{S^{t,e}})}\\&=\frac{\sum_{i=1}^{N}\sum_{\overline{S^{t,e}}\in\overline{S^e},m\in M}(1-p_i^d)p_i^g\eta_{i,m}^t(\overline{S^{t,e}})}{\sum_{i=1}^{N}\sum_{\overline{S^{t,e}}\in\overline{S^e},m\in M}(1-p_i^d)p_i^g\eta_{i,m}^t(\overline{S^{t,e}})+\sum_{i=1}^{N}\sum_{\overline{S^{t,e}}\in\overline{S^e},m\in M}p_i^d(1-p_i^g)\eta_{i,m}^t(\overline{S^{t,e}})}\end{aligned} \tag{9-29}$$

结合式(9-28)可得随机演化联盟博弈的 Q 函数为

$$\begin{aligned}Q(S^t,a^t,\overline{a^t},a_h^t)=&C(S^t,a^t,\overline{a^t},a_h^t)\eta_t(\overline{S^{t,e}})\\&+\gamma^t\sum_{S^{t,d},S^{t+1,d}\in S^d,\overline{S^{t,e}},\overline{S^{t+1,e}}\in\overline{S^e}}P(S^{t+1,d},\overline{S^{t+1,e}}\mid S^{t,d},\overline{S^{t,e}},a^t,\overline{a^t},a_h^t)\\&\times V(S^{t+1,d},\eta_{t+1}(\overline{S^{t+1,e}}),\pi^*)\end{aligned} \tag{9-30}$$

式中，$\eta_t(\overline{S^{t,e}})$为在时刻 t 的演化联盟的可靠性更新函数值；$V(S^{t+1,d},\eta_{t+1}(\overline{S^{t+1,e}}),\pi^*)$为随机演化联盟博弈状态更新的值函数，通过估计值函数获得随机演化联盟博弈的最优策略 π^*。

令策略 π 表示一个从状态到行动的映射 $\pi:S\to A$。在攻击者的攻击下，随机演化联盟博弈状态更新的值函数表示为

$$V(S^{t+1,d},\eta_{t+1}(\overline{S^{t+1,e}}),\pi^*)=\max_{\pi(a^t,\overline{a^t})}\min_{\pi(a_h^t)}\sum_{a^t\in a,\overline{a^t}\in\overline{a},S^t\in S}Q(S^t,a^t,\overline{a^t},a_h^t)\pi(a^t,\overline{a^t}) \tag{9-31}$$

式中，$\pi(a,\overline{a})$为随机演化联盟博弈参与者的防御策略；$\pi(a_h^t)$为攻击者采取的攻击策略。从式(9-31)可看出，如果求得随机演化联盟博弈 $Q(S^t,a^t,\overline{a^t},a_h^t)$的值最大时的优化策略 π_d^*，

就可以通过观察在状态 S^t 下最大化演化联盟的行动收益 $V_t(S^{t+1,d},\eta_{t+1}(\overline{S^{t+1,e}}),a)$ 来获得优化策略 π_e^*，从而得出随机演化联盟博弈的最优策略 $\pi^*=\pi_d^*\pi_e^*$。Q－learning 实现了 Q 函数更新过程，联盟参与者首先初始化 $Q(S^t,a^t,\overline{a^t},a_h^t)$，然后更新 Q 函数为

$$Q_{t+1}(S^t,a^t,\overline{a^t},a_h^t)=(1-\alpha_t)Q_t(S^t,a^t,\overline{a^t},a_h^t)+\alpha_t[C(S^t,a^t,\overline{a^t},a_h^t)\eta_t(s^t)+\gamma^t\max_{a\in\overline{a}}V_t(S^{t+1,d},\eta_{t+1}(\overline{S^{t+1,e}}),a)] \tag{9-32}$$

式中，$\alpha_t\in[0,1]$为学习率。

基于 Q－learning 的随机演化联盟博弈算法包括两个方面：一方面是联盟参与者对于动态变化的攻击策略的自适应防御；另一方面是联盟参与者对于联盟服务组合攻击策略的动态演化联盟博弈。详细过程如算法 9-2 所示。

算法 9-2 优化虚拟传感云服务可靠性和 QoS 的随机演化联盟博弈算法。

输入：博弈状态 S，行动集合 A，收益函数 δ。

输出：纳什均衡策略集 π^*，演化联盟均衡策略集 M^*。

1. 初始化 $t=0$。
2. FOREACH $S^t\in S$ AND $a^t\in a,\overline{a^t}\in\overline{a}$ AND $i\in\{1,\cdots,\overline{N}\}$
3. 　$Q_t^i(S^t,a^t,\overline{a^t},a_h^t)=1$。
4. 　$V_t^i(S^t,\eta_t(s^t),\pi^*)=1$。
5. ENDFOR
6. LOOP
7. 　攻击者随机采取攻击行动 a_h^t，联盟参与者随机采取防御行动 $a_i^t,\overline{a^t}$，使得联盟节点状态由 S^t 变为 S^{t+1}，联盟成员得到的收益为 $C(S^t,a^t,\overline{a^t},a_h^t)$。
8. 　FOR $i=1$ TO $\overline{N}$
9. 　　根据式(9-29)计算演化联盟的可靠性更新函数值 $\eta_{t+1}(\overline{S^{t+1,e}})$。
10. 　　更新优化策略和状态 $V_t^i(S^{t+1,d},\eta_{t+1}(\overline{S^{t+1,e}}),\pi^*)$。
11. 　　更新函数 $Q_t^i(S^t,a^t,\overline{a^t},a_h^t)$。
12. 　ENDFOR
13. 　联盟参与者 i 分别计算联盟 M_k 和 M_v 期望的收益 $\delta_i(M_k,\eta_{M_k}^i)$和 $\delta_i(M_v,\eta_{M_v}^i)$。
14. 　IF $(\delta_i(M_k,\eta_{M_k}^i)>\tilde{\omega}_i)$ AND $(\eta_j^{c,k}=1)$ AND $\delta_i(M_v,\eta_{M_v}^i)>\tilde{\omega}_i)$ AND $(\eta_j^{c,v}=1))$
　//$\tilde{\omega}_i$ 表示期望收益的阈值，$\eta_{M_k}^i$ 表示虚拟传感节点 i 期望与联盟 M_k 中的邻居
　//节点 j 合作，$\eta_j^{c,k}=1$ 表示虚拟传感节点 i 能够与联盟 M_k 中的邻居节点 j
　//合作，$\eta_{M_v}^i$ 表示虚拟传感节点 i 期望与联盟 M_v 中的邻居节点 j 合作，$\eta_j^{c,v}=1$
　//表示虚拟传感节点 i 能够与联盟 M_v 中的邻居节点 j 合作。
15. 　　IF $\delta_i(M_k\cup\{j\},\eta_{M_k}^i)>\delta_i(M_v\cup\{j\},\eta_{M_v}^i)$
16. 　　　虚拟传感节点 i 期望在联盟 M_k 中选择邻居节点 j 加入联盟 M^t。
17. 　　ENDIF
18. 　　IF $M_v\rhd_i M_k$

联盟参与者 i 在联盟 M_k 中选择邻居节点 j 加入联盟 M^t

$M^{t+1}=M^t\cup\{j\}\cup M_k\setminus\{j\}$。

ELSE

$M^{t+1}=M^t$。

ENDIF

ENDIF

$t=t+1$。

UNTIL $Q_{t+1}^i(S^t,a^t,\overline{a^t},a_h^t)=Q_t^i(S^t,a^t,\overline{a^t},a_h^t)$

FOREACH $S^t\in S$

$\pi^*=Q_{t+1}^i(S^t,a^t,\overline{a^t},a_h^t)$。

$M^*=M_1^*\cup M_2^*\cdots\cup M_M^*$。

ENDFOR

RETURN π^*,M^*。

9.6 随机演化联盟均衡学习策略

9.6.1 基于 Shapley 值的多重收益分配

定义 9-8 随机演化联盟均衡策略为

$$C_i(\pi_i^*,\pi_{-i}^*)\geqslant C_i(\pi_i,\pi_{-i}^*) \tag{9-33}$$

式中,$C_i(\cdot)$为虚拟服务组合联盟获得的收益;π_i 为联盟参与者 i 的策略,π_i^* 表示联盟参与者 i 的纳什均衡策略;π_{-i}^*为除联盟参与者 i 外的所有联盟参与者的纳什均衡策略。随机演化联盟博弈达到平衡后,收益的平均分配决定了在时刻 $t+1$ 联盟参与者之间合作的动力。Shapley 值使用向量$\boldsymbol{\varphi}(c)=(\varphi_1(c),\cdots,\varphi_{\bar{f}}(c))$分配收益,其中 $\varphi_i(c)$表示每个联盟参与者 i 的平均收益分配。令在联盟 M 中,向内提供服务的虚拟传感节点总数为$|M^{in}|$,向外提供服务的虚拟传感节点总数为$|M^{ex}|$。

由此可得,虚拟传感节点 j 向内提供服务时,获得 Shapley 值的平均收益分配概率为

$$\varphi_j^{in}(c)=\sum_{M^{in}\subseteq(M-M^{ex})\setminus\{j\}}\frac{|M^{in}|!(|M|-|M^{in}|-|M^{ex}|-1)!}{|M|!}[c(M^{in}\cup\{j\})-c(M^{in})] \tag{9-34}$$

式中,$c(M^{in}\cup\{j\})-c(M^{in})$为博弈参与者 j 加入内部服务组合联盟 M^{in}时提供的共享容量;$c(M^{in})$为内部服务组合联盟的容量;$c(M^{in}\cup\{j\})$为博弈参与者 j 加入内部服务组合联盟 M^{in}后内部服务组合联盟的容量。

类似地,虚拟传感节点 j 向外提供服务时,获得 Shapley 值的平均收益分配概率为

$$\varphi_j^{ex}(c)=\sum_{M^{ex}\subseteq\{M-M^{in}\}\setminus\{j\}}\frac{|M^{ex}|!(|M|-|M^{in}|-|M^{ex}|-1)!}{|M|!}[c(M^{ex}\cup\{j\})-c(M^{ex})] \tag{9-35}$$

式中,$c(M^{ex}\cup\{j\})-c(M^{ex})$为博弈参与者 j 加入外部服务组合联盟 M^{ex}时提供的共享容量;$c(M^{ex})$为外部服务组合联盟的容量;$c(M^{ex}\cup\{j\})$为博弈参与者 j 加入外部服务组合联盟

M^{ex}后，外部服务组合联盟的容量。

由此可得，随机演化联盟博弈中虚拟传感节点 j 的 Shapley 值平均收益分配的容量为

$$c_t^j(S^t,a^t,\varphi_j(c)) = \varphi_j^{in}(c)[c_{t[0]}^{in}(S^t,a^t)+c_{t[1]}^{in}(S^t,a^t)]+\varphi_j^{ex}(c)[c_{t[0]}^{ex}(S^t,a^t)+c_{t[1]}^{ex}(S^t,a^t)] \tag{9-36}$$

式中，$c_{t[0]}^{in}$ 和 $c_{t[1]}^{in}$ 分别为内部服务不合作和合作获得的收益分配；$c_{t[0]}^{ex}$ 和 $c_{t[1]}^{ex}$ 分别为外部服务不合作和合作获得的收益分配。

Shapley 值用于随机演化联盟博弈的合作收益分配，主要有以下特性：

(1) $\sum_{j\in M} c_t^j(S^t,a^t,\varphi_j(c)) > c(M)$，其中 $c(M)$ 表示联盟 M 的平均容量，表明随机演化联盟博弈总的合作收益值变大。

(2) 如果 $c(M\cup\{i\})=c(M\cup\{j\})$，则 $c_t^i(S^t,a^t,\varphi_i(c))=c_t^j(S^t,a^t,\varphi_j(c))$，表明虚拟传感节点 i 和 j 具有相同的收益分配，也即获得了保证其可靠性的容量，使得联盟 $M\cup\{i\}$ 和 $M\cup\{j\}$ 具有相同的安全级别。

(3) 如果 $c(M)=c(M\cup\{i\})$，则 $c_t^i(S^t,a^t,\varphi_i(c))=0$，这表明虚拟传感节点 i 不能提供共享容量，因此，在本次博弈中获得的收益为零，这使得服务组合的安全级别降低。

(4) 如果 $c(M_1\cup M_2)\geqslant c(M_1)+c(M_2)$，并且 $M_1\cap M_2=\varnothing$，这表明联盟 M_1 和 M_2 合作的收益大于不合作的收益，使得合作后服务组合的安全级别升高。

9.6.2 随机演化联盟的收益估计

随机演化联盟的收益估计考虑两方面，即虚拟传感节点组成的服务组合既向私有云提供内部服务，又向公有云提供外部服务。因此，在同等条件下，虚拟传感节点要在保证私有云的内部服务正常运行的情况下再向外部提供服务，即内部服务的收益大于外部服务的收益。博弈参与者 j 在内部服务组合联盟中的收益估计为

$$\tilde{c}_{m,t+1}^{j,in} = \tilde{c}_{m,t}^{j,in} + \sigma_{j,t}^{in}(\mathbf{1}_{\{a_{j,t}=M^{in}\}} + \mathbf{1}_{\{C(S^t,a^t,a_h^t)\geqslant\delta_t(M^{in},\eta_{M^{in}}^i)\}})(c_{m,t}^{j,in} - \tilde{c}_{m,t}^{j,in}) \tag{9-37}$$

式中，$\tilde{c}_{m,t}^{j,in}$ 为在时刻 t 内部服务组合联盟 m 中博弈参与者 j 期望的收益；$\sigma_{j,t}^{in}$ 为在时刻 t 博弈参与者 j 在内部服务组合联盟的学习速率；$\mathbf{1}_{\{a_{j,t}=M^{in}\}}$ 为指示函数，表明在时刻 t 联盟参与者 j 选择内部服务组合联盟 M^{in} 加入时值为 1，否则为 0；$\mathbf{1}_{\{C(S^t,a^t,a_h^t)\geqslant\delta_t(M^{in},\eta_{M^{in}}^i)\}}$ 为指示函数，表明攻击者对内部服务组合联盟 M^{in} 采用的行动为 a_h^t，联盟参与者采取行动为 a^t，获得的收益为 $C(S^t,a^t,a_h^t)\geqslant\delta_t(M^{in},\eta_{M^{in}}^i)$ 时，函数值为 1，否则为 0；$c_{m,t}^{j,in}$ 为在时刻 t 联盟参与者 j 选择内部服务组合联盟 m 后观察到的收益。

同理，博弈参与者 j 在外部服务组合联盟中的收益估计为

$$\tilde{c}_{m,t+1}^{j,ex} = \tilde{c}_{m,t}^{j,ex} + \sigma_{j,t}^{ex}(\mathbf{1}_{\{a_{j,t}=M^{ex}\}} + \mathbf{1}_{\{C(S_t,a^t,a_h^t)\geqslant\delta_t(M^{ex},\eta_{M^{ex}}^i)\}})(c_{m,t}^{j,ex} - \tilde{c}_{m,t}^{j,ex}) \tag{9-38}$$

式中，$\tilde{c}_{m,t}^{j,ex}$ 为在时刻 t 外部服务组合联盟 m 中博弈参与者 j 期望的收益；$\sigma_{j,t}^{ex}$ 为在时刻 t 博弈参与者 j 在外部服务组合联盟的学习速率；$\mathbf{1}_{\{a_{j,t}=M^{ex}\}}$ 为指示函数，表明在时刻 t 联盟参与者 j 选择外部服务组合联盟 M^{ex} 加入时值为 1，否则为 0；$\mathbf{1}_{\{C(S^t,a^t,a_h^t)\geqslant\delta_t(M^{ex},\eta_{M^{ex}}^i)\}}$ 为指示函数，表明攻击者对外部服务组合联盟 M^{ex} 采用的行动为 a_h^t，联盟参与者采取行动为 a^t，获得的收益为 $C(S^t,a^t,a_h^t)\geqslant\delta_t(M^{ex},\eta_{M^{ex}}^i)$ 时，函数值为 1，否则为 0；$c_{m,t}^{j,ex}$ 为在时刻 t 联盟参与者 j 选择外部服务组合联盟 m 后观察到的收益。值得注意的是，在一般情况下，$\tilde{c}_{m,t+1}^{j,in}\geqslant\tilde{c}_{m,t+1}^{j,ex}$。

9.6.3 随机演化联盟的策略学习

要使得整个联盟获得最大收益,新加入联盟的参与者应持续地学习其他联盟成员的策略,达到所有联盟成员联合防御的目的。因此,策略学习的过程是联盟成员复制其他联盟成员行动的过程。这表明一个新加入联盟的参与者将以较高的概率复制联盟内成功防御的参与者的行动。这个过程的实现主要通过观察实施成功防御的参与者的收益来获得,一个参与者防御后的收益越高说明其防御越成功,其策略及行动被复制的概率就越大。当参与者发现一个参与者防御后的收益很低,它将采取非理性的方式来复制策略及行动。策略学习方程表示为

$$\theta^{j}_{m,t+1} = \frac{\theta^{j}_{M,t}\eta^{j}_{M,t}\ (1+\xi_t)^{p_{i\to j}}}{\sum\limits_{M'}\theta^{j}_{M',t}\eta^{j}_{M',t}\ (1+\xi_t)^{p_{i\to j}}} \cdot \Delta_j(t,t+1) \tag{9-39}$$

式中,$\theta^{j}_{M,t}$为若在时刻t 虚拟传感节点j 被选中加入到联盟M,则其值为1,否则为0;$\eta^{j}_{M,t}$为若在时刻t 虚拟传感节点j 处于可靠状态,则其值为1,否则为0;ξ_t 为在时刻t 博弈参与者j 的学习速率。$p_{i\to j}$的定义详见式(9-4),表示复制概率。联盟参与者j 变化其策略的概率为

$$\Delta_j(t,t+1) = \begin{cases} \chi, & c_j(t+1) > c_j(t) \\ 1-\chi, & \text{其他} \end{cases} \tag{9-40}$$

在式(9-40)中,如果联盟参与者j 发现在时刻t 复制其他联盟策略后,在时刻$t+1$ 的收益大于时刻t 的收益,即$c_j(t+1)>c_j(t)$表示联盟参与者j 将以理性的概率χ 变化其策略。如果出现$c_j(t+1)\leqslant c_j(t)$时,联盟参与者j 将以非理性的概率$1-\chi$ 变化其策略。

从式(9-23)、式(9-39)、式(9-40)可以看出,随机演化联盟博弈的稳定状态决定于马尔可夫链的传递概率的稳定性。当联盟参与者以很小的非理性策略变化时,随机演化联盟博弈达到稳定平衡状态。随机演化联盟博弈的稳定状态吸收了随机博弈和演化联盟博弈的稳定状态,它与马尔可夫链的稳定状态相一致。当随机演化联盟博弈达到稳定状态时,容量收益不能够再分配,此时,可以表示随机演化联盟博弈稳定状态为$\Omega^*(S^*,P^*,\pi^*,\delta^*,M^*)$。其中,$S^*$表示随机演化联盟稳定的状态;$P^*$表示稳定状态的传递概率;$\pi^*$表示稳定策略;$\delta^*$表示稳定状态的收益;$M^*$表示稳定联盟结构。

9.7 实验

本章利用 MATLAB R2010a 仿真了已提出的模型和算法。在仿真实验中,首先评估了BA 模型生成可靠联盟时参数α 和τ 对算法性能的影响。对于基于 BA 模型的可靠联盟的形成,设置无标度网络初始节点个数为 20,每次选择新的可靠节点加入联盟后生成的边数为 2,生成可靠联盟后的网络规模为 100。初始网络节点都为孤立节点,生成的可靠联盟中各个节点的容量分布如图 9-3 所示。从图 9-3 中可以看出,随着节点数的增加,联盟中的每个参与者共享的容量数在减少。当节点数为 20 时,联盟中每个参与者共享的容量数达到最大。图 9-4 给出了联盟中节点的容量分布,其平均容量分配值约为 3.14。从图 9-4 中可以看出,最终形成联盟的容量数为 3 的节点达到 43%。没有加入联盟的节点达 17%。通过

BA 模型使得不可靠的节点未加入联盟，因此，减少了形成联盟的节点数，使得随机演化联盟防御阶段的演化稳定时间缩短。

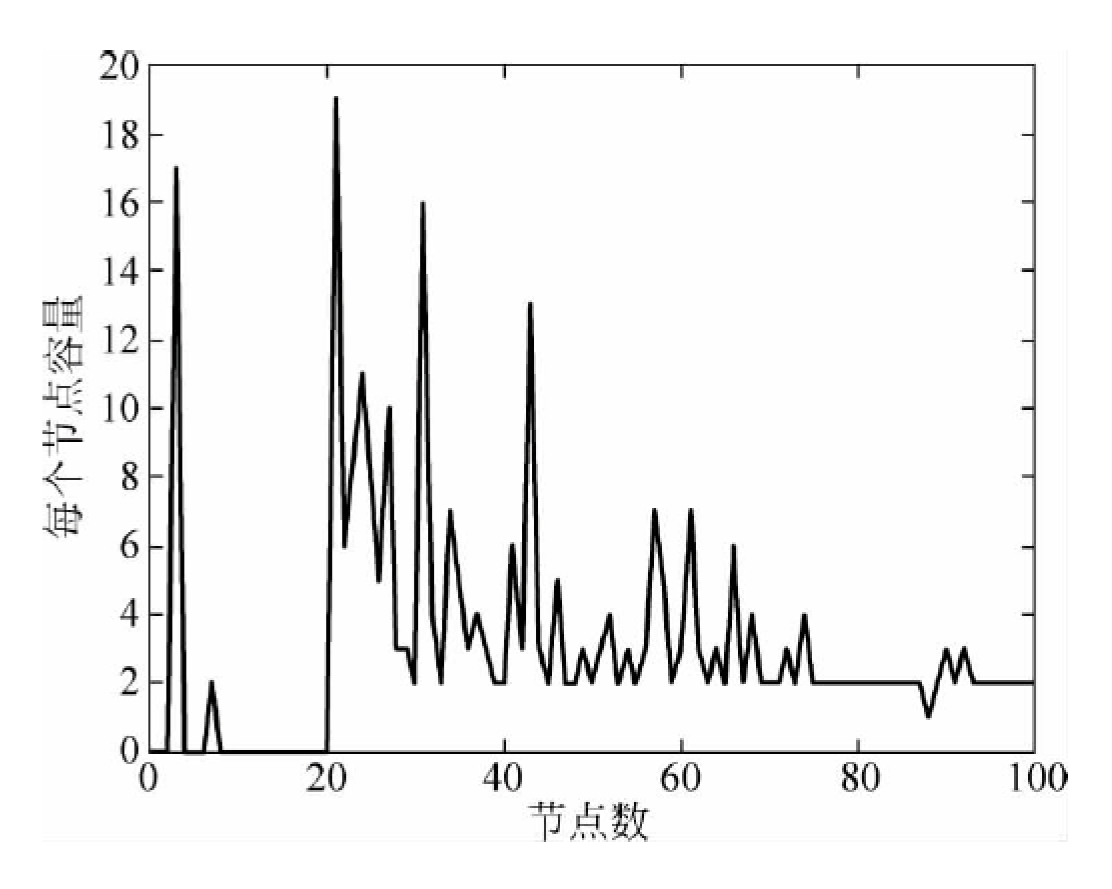


图 9-3　联盟中每个节点的容量分布

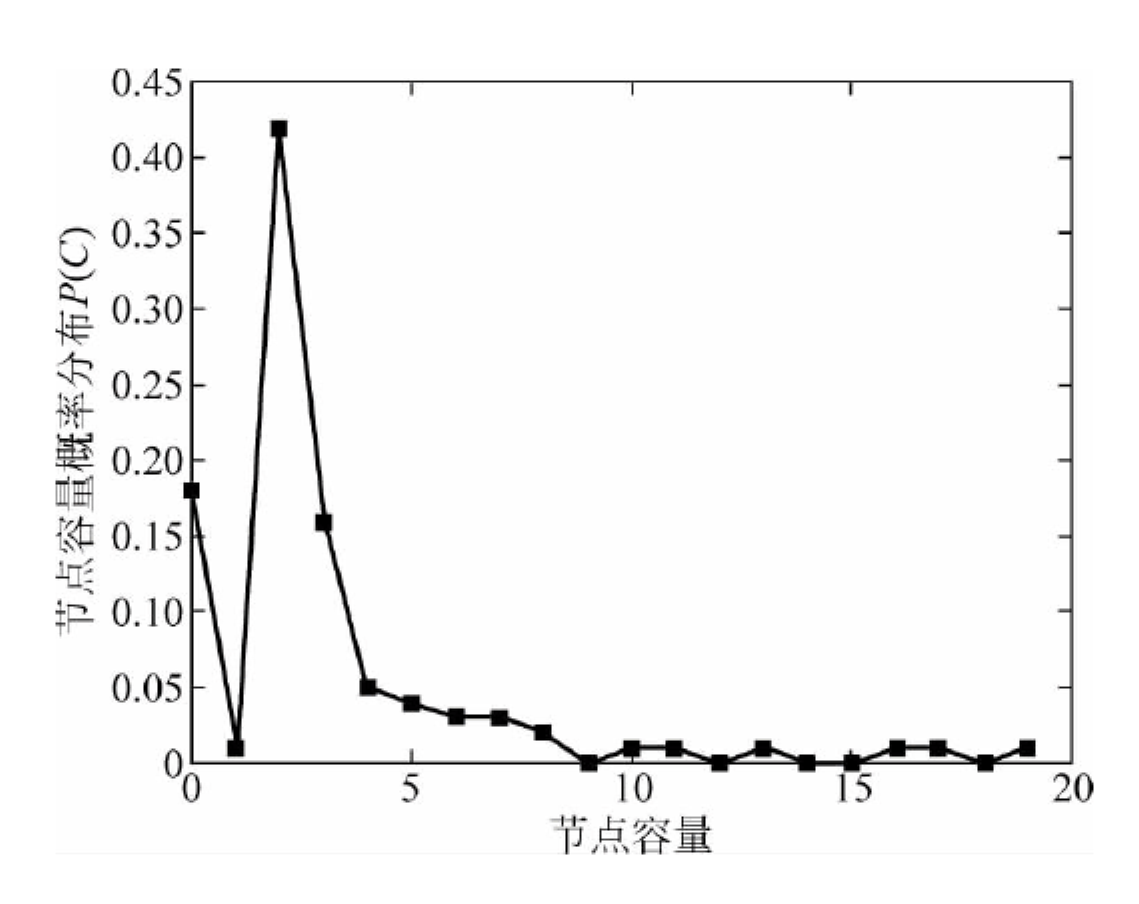


图 9-4　联盟中节点容量的概率分布

图 9-5 给出了在 $\alpha\leqslant 0$ 的情况下，参数 γ、α 值对应联盟节点间合作速率的影响。从图 9-5 可以看出，随着 γ 的减少，合作速率 υ 降低，这表明越小的贴现率，合作的机会越小，收益越小，虚拟传感节点的可靠性越差。图 9-6 显示了在 $\alpha>0$ 的情况下，参数 γ、α 值对联盟节点间合作速率的影响。从图 9-6 可以看出，随着 γ 的增大，合作速率 υ 升高。这表明越大的贴现率，合作的机会越多，收益越高，虚拟传感节点的可靠性越高。

图 9-7 给出了攻击者和防御者的学习曲线。面对攻击时，防御者开始有最小的收益，随后，防御者观察到攻击者的策略和行动后，防御者变化其策略，通过 Q 值和可靠性更新，联盟参与者选择可靠的虚拟传感节点并通过共享容量来提高其收益，使得收益达到局部最大值。经过连续多次学习和迭代后，攻击者的收益达到局部最小。防御者通过 Q 函数累积其收益后，使其收益达到最大，而攻击者累积的收益达到最小。图 9-8 给出了不同博弈的平均收益，随机演化联盟博弈(SECG)由于使用了 BA 模型形成演化联盟，开始其收益升高到约为 0.5，接着逐渐升高到最大值，约为 0.68，随后面对攻击时，防御者和攻击者开始博弈，经过 6 步迭代后达到均衡，此时，SECG 收益约为 0.57。随机博弈(SG)模型虽然没有使用联

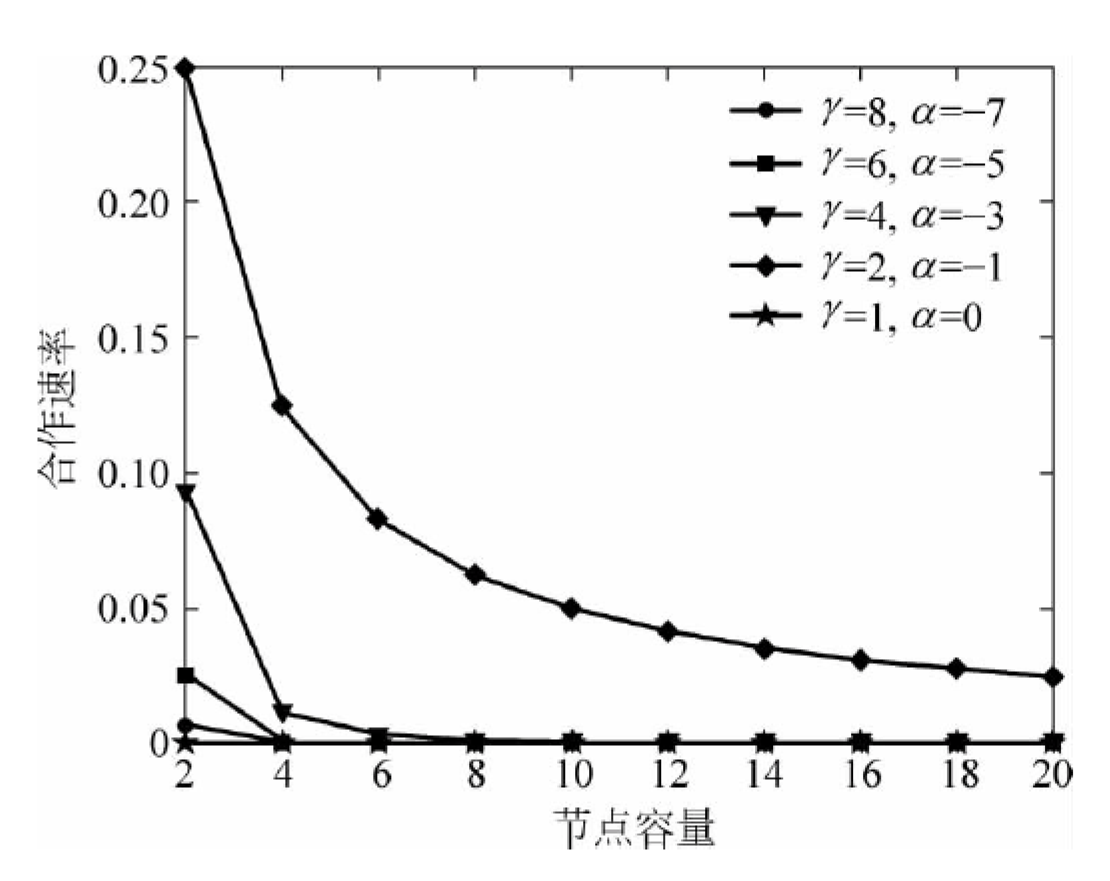


图 9-5 打折因子的大小和 $\alpha\leqslant 0$ 对合作速率的影响

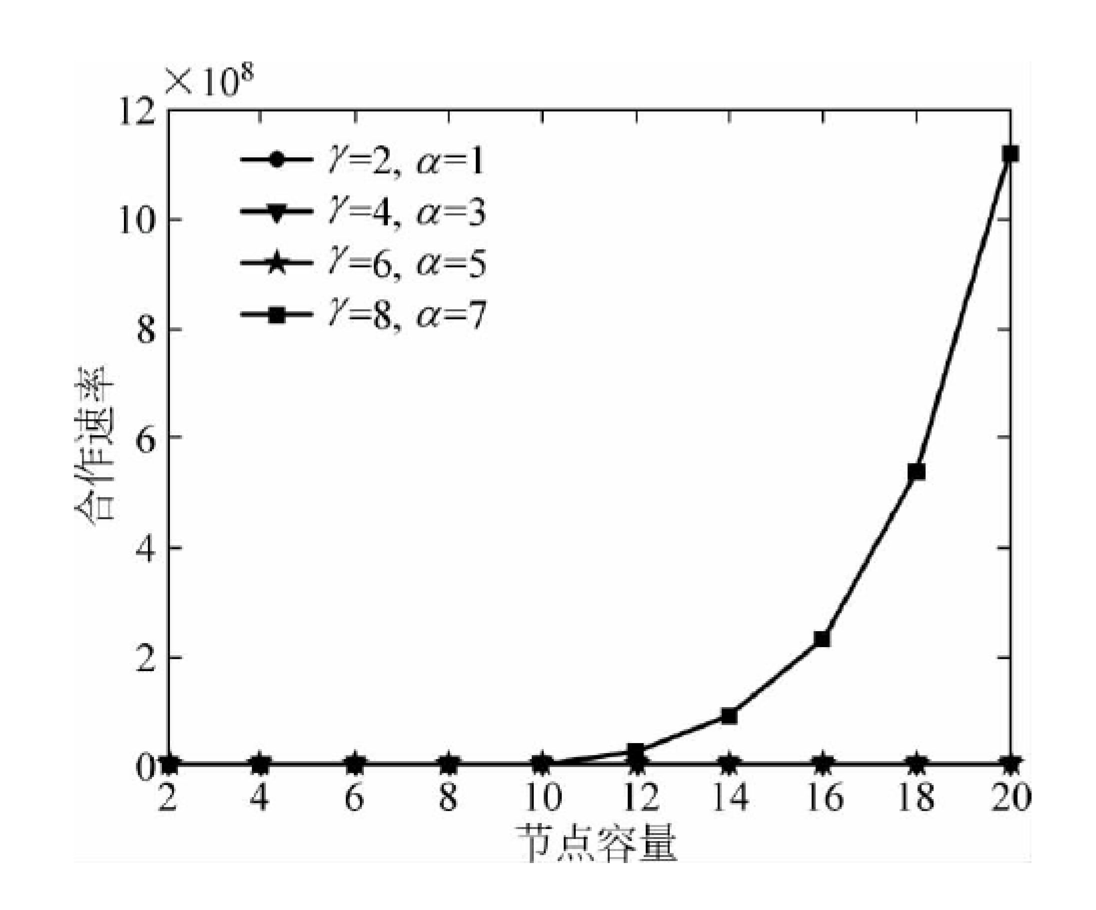


图 9-6 打折因子的大小和 $\alpha>0$ 对合作速率的影响

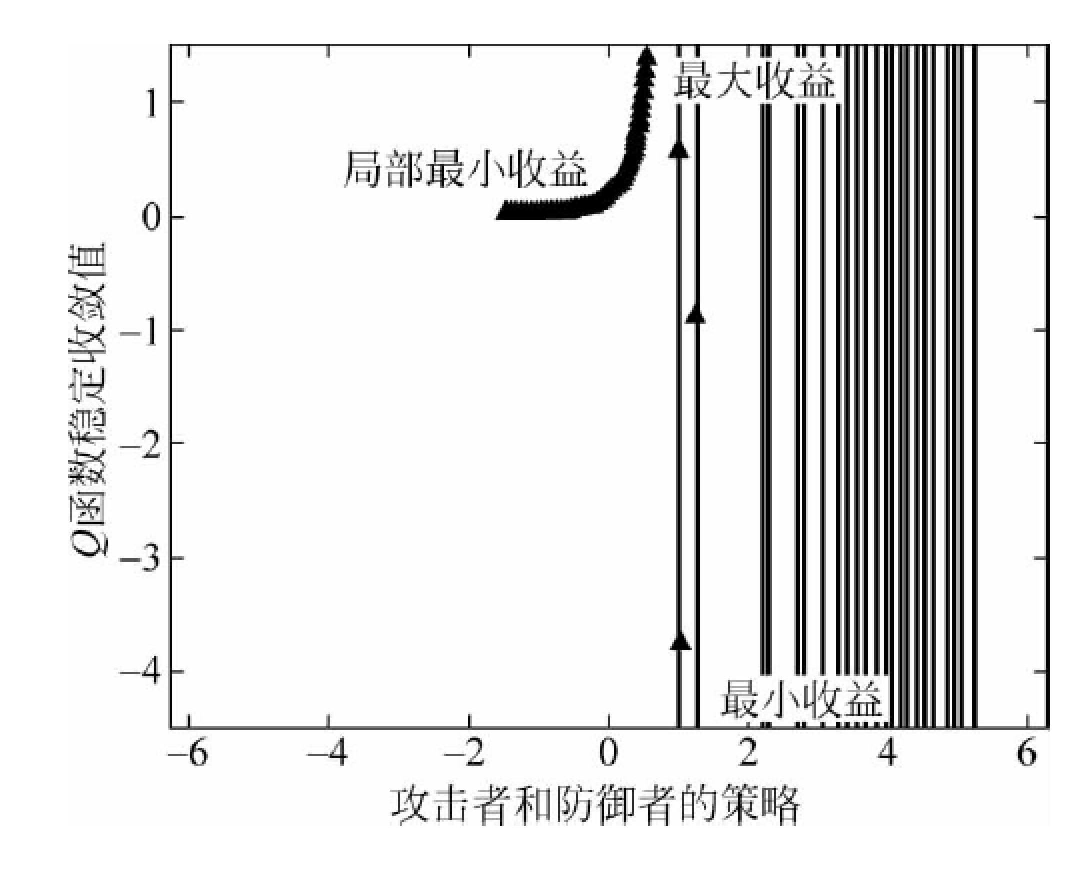


图 9-7 攻击者和防御者的策略和稳定性学习曲线

盟博弈，但是由于随机的策略选择能有效地防御攻击，开始收益约为 0.4，经过 5 步迭代达到稳定状态。演化联盟博弈(ECG)的平均收益约为 0.2，由于只靠联盟的演化来防御，不能自适应地变化防御策略，这使得当联盟攻击者的策略变化时，ECG 的收益减少，联盟成员开

始选择新的联盟加入，经过3步迭代后，联盟达到收益的最高点并达到均衡状态。若此时再受到攻击，联盟开始下一轮的演化，同样经过3步迭代后达到均衡状态。由于SECG包含联盟演化的过程，SG比SECG先到达均衡。

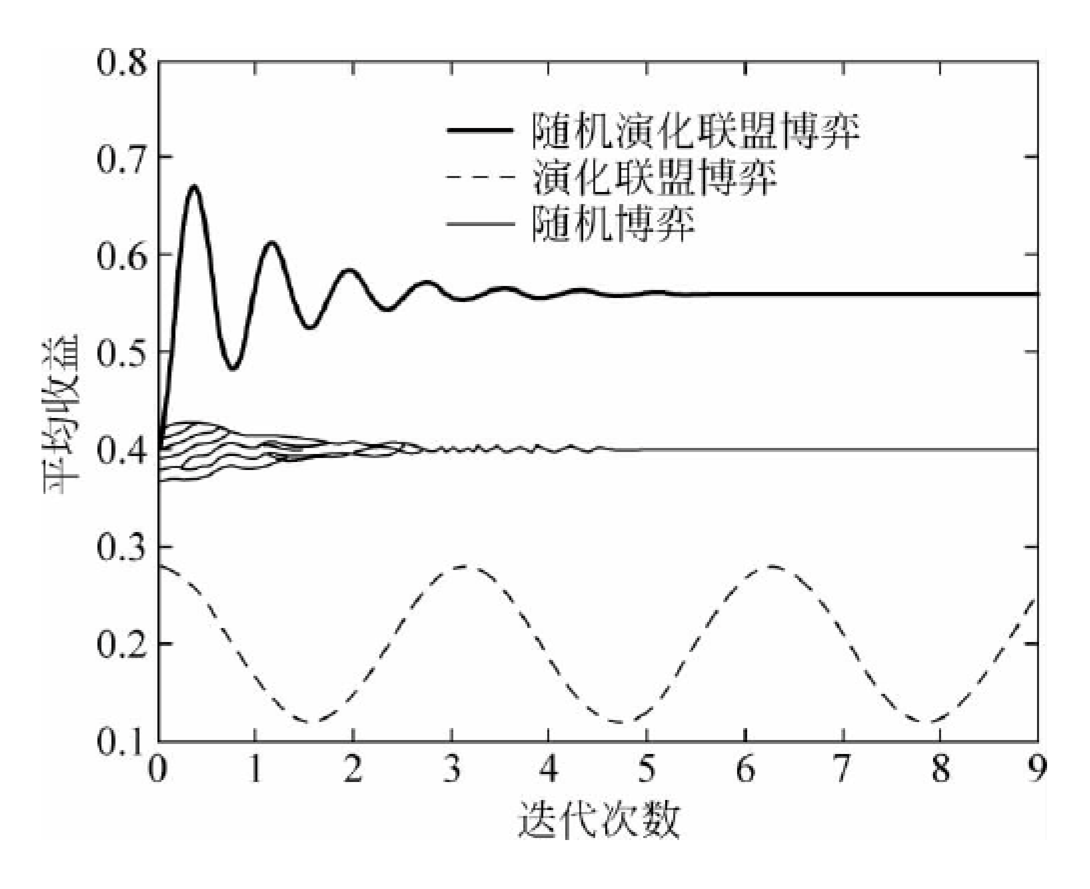


图9-8　不同博弈的平均收益

图9-9给出了不同策略的总贴现收益，其中SECG在优化联盟当前结构时，考虑了将来的收益，实现了很高的贴现收益，SECG比SG高86%，比固定策略高出约2倍多。因此，使用SECG能动态适应攻击者策略的变化，有效地提高虚拟传感云服务的可靠性和服务质量。图9-10给出了不同博弈中联盟结构大小对于网络通信成本的影响。其通信成本计算式为

$$C_{\mathrm{tr}} = \sum_{i=1}^{\bar{M}} \Upsilon_i \zeta_i (d_r^i + d_p^i) \tag{9-41}$$

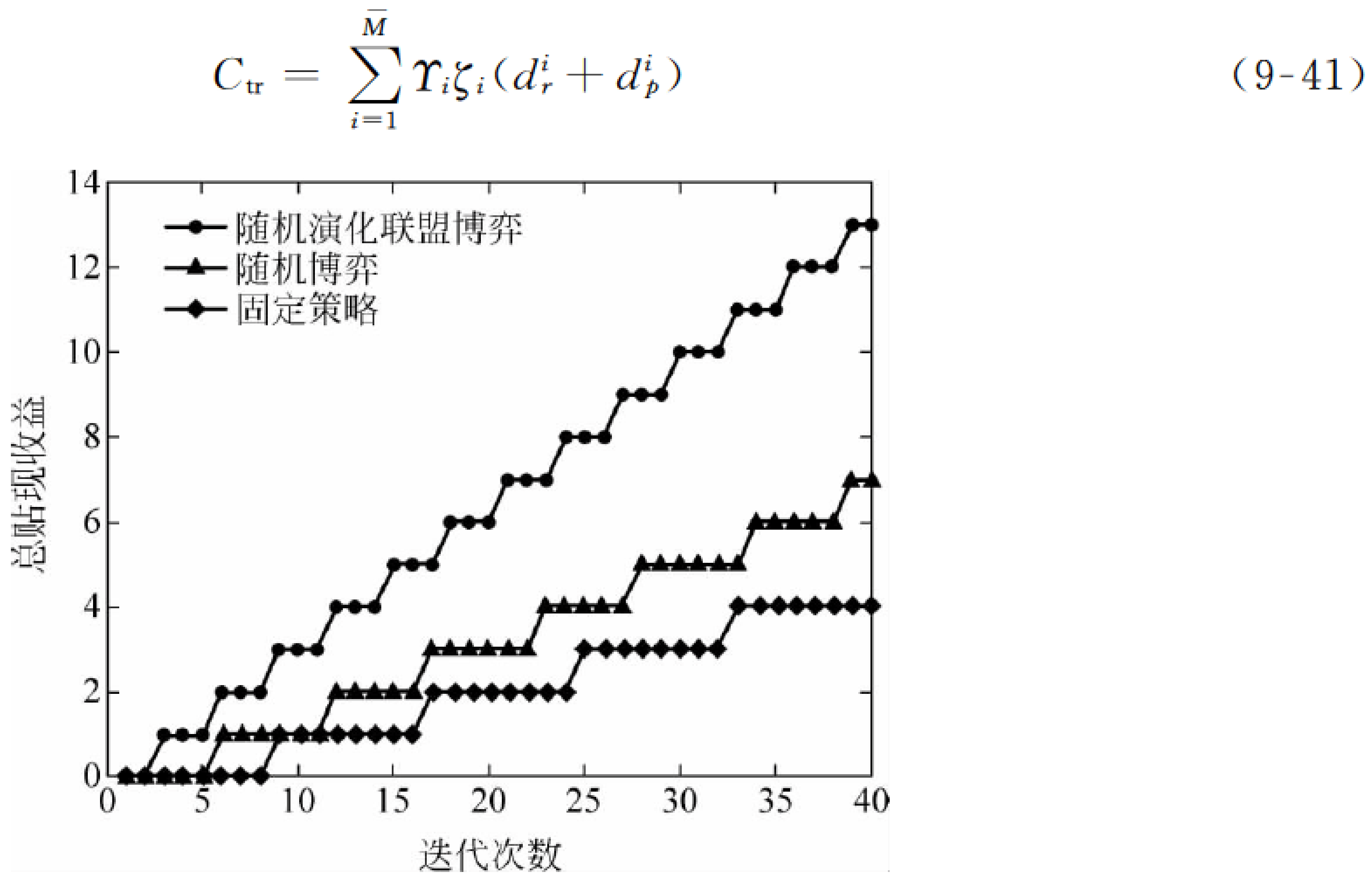


图9-9　不同策略达到稳定状态时的贴现收益

式中，Υ_i为虚拟传感云服务组合联盟i被请求消息的次数；ζ_i为消息在虚拟传感云服务组合联盟i中传递的跳数；d_r^i为虚拟传感云服务组合联盟i请求消息的数据量；d_p^i为虚拟传感云服务组合联盟i接收消息的数据量。从图9-10中可以看出，随着联盟结构规模的增加，不同博弈的通信成本呈现上升趋势，其中演化联盟博弈ECG的增长快于SECG和SG。这是由于在ECG中没有预先形成联盟，因此，联盟参与者必须先和所有的虚拟传感节点通

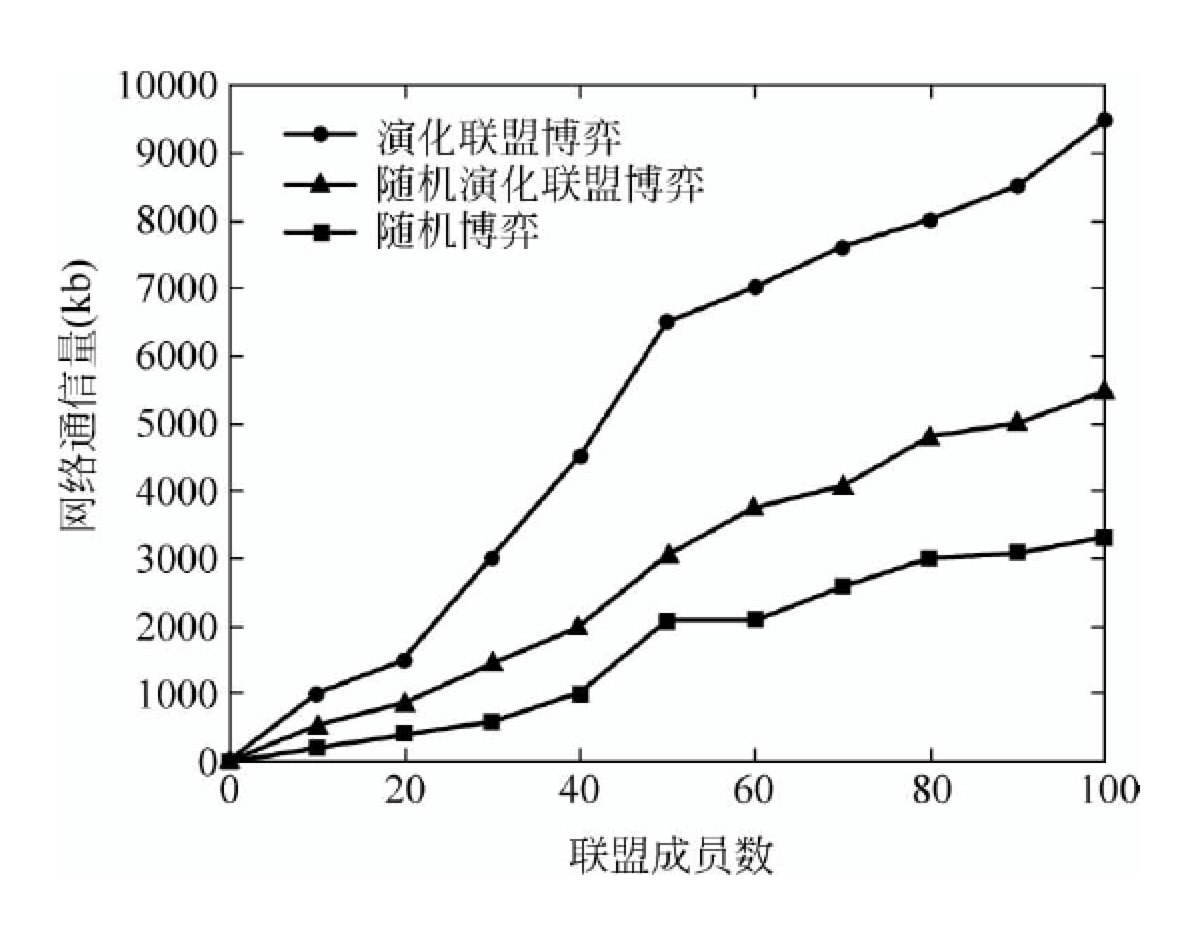


图 9-10 不同博弈联盟结构对网络通信量的影响

信形成联盟，再演化形成可靠的联盟。SECG 先通过 BA 模型形成了可靠的联盟，在随机演化博弈过程中可以不必和所有的虚拟传感节点通信，从而减少通信量。SG 不需要形成联盟，只需与入侵检测系统传递消息，因此其通信成本最小。对于联盟大小为 80 的联盟而言，ECG 比 SECG 的通信成本高约 67%，而比 SG 高出约 1 倍多。SECG 比 ECG 的通信成本小约 40%。对于联盟大小为 40 的联盟而言，ECG、SECG 博弈策略达到均衡后，联盟处于稳定状态，此时通信量不再变化，当稳定状态概率发生变化时，意味着联盟博弈转换为下一个状态，此时通信量又开始增加，但不会超过前一个状态的通信量，这是由于稳定状态概率的变化是由攻击者的攻击引起，此时的联盟参与者作为防御者只需使得可靠性较差、受攻击的联盟进行演化，而不会与联盟前一状态的所有成员通信。

9.8 小结

为了提高虚拟传感云服务的可靠性和服务质量，本章分析了虚拟传感云服务系统动态的容量和变化的攻击者策略，模型化攻击者和防御者之间的交互过程为一个随机演化联盟博弈。在模型中，联盟参与者根据对系统容量状况和攻击行动的观察，使用学习算法自适应攻击者策略的变化，并根据它们的偏好理性地形成联盟来最大化个体的收益。仿真结果显示 SECG 模型能有效减少网络通信量和获得最大收益，并且能较快达到稳定状态。与 ECG 和固定策略相比，SECG 模型能够获得较高的收益和性能。使用 Q 学习算法，通过多次迭代，使得防御者的累积收益达到最大，攻击者的累积收益达到最小。此外，本章还阐述了联盟的可靠性更新和基于 Shapley 值的多重收益分配机制，在 SECG 模型中使用这个机制，可以不断更新虚拟传感节点的可靠状态和累积收益，加快联盟的稳定和激励下一轮的合作。

第 10 章

基于演化博弈的传感器节点保密率自适应调节研究

本章通过扩展经典窃听信道模型，针对聚簇无线传感器网络提出了传感器节点和其对应簇头节点之间的保密率计算方法，构建了一个非合作保密率博弈模型，以反映传感器节点之间的交互关系。利用演化博弈思想，建立了传感器节点自适应选择发射功率的机制，提出了传感器节点自适应调节保密率的算法。实验结果表明，提出的方法能自适应地调节传感器节点的保密率，为保证无线传感器网络数据的保密性提供了新途径。

10.1 引言

传感器节点资源的有限性决定了保障无线传感器网络的安全具有很大的挑战。国内外学者为防御无线传感器网络中存在的威胁和漏洞，提出了加密机制、攻击检测、安全路由等多种方法[11]。与传统保障网络通信安全方法不同，物理层安全利用无线信道的物理特性保障无线通信的安全[442, 443]。由于这种方法不需要添加额外组件，而是利用物理层的基础能力实现网络通信安全，所以非常适合资源有限的无线传感器网络。

物理层安全技术的实质是在防止窃听者获得通信数据的同时，最大化源节点到目标节点的可靠通信率(保密率)。其中，最大化的保密率被称为保密容量。Wyner 在其开创性的工作[444]中介绍了经典窃听模型，并说明了即使不使用密钥也能实现节点间的绝对通信安全。随后，国内外学者对高斯窃听信道[445]、广播信道[446]、中继信道[447]、干扰信道[448]、MISO(Multi-Input Single-Output，多输入单输出)信道[449, 450]、MIMO(Multi-Input Multi-Output，多输入多输出)信道[451, 452]等的保密容量进行了系统的基础研究。

在各种信道保密容量研究基础上，当前的一个焦点问题是如何在相同的网络环境下提高物理层的安全性。然而，现有文献在提高物理层安全性的同时，未考虑对网络通信的影响。实际上，在无线传感器网络数据通信过程中，发射功率是影响传感器节点保密率的关键因素。由于传感器节点的自私性，每个传感器节点都希望增强自身的发射功率，从而最大化自己的保密率。但这种增强自身发射功率的做法干扰了其他传感器节点的通信，同时大大消耗了自身能量。因此，仅考虑最大化传感器节点保密率，对整个网络的效用而言不一定是最优的选择。

本章以最大化网络效用为目标，利用演化博弈方法实现传感器节点保密率的自适应调

节。首先,为适应无线传感器网络环境扩展了经典保密率计算公式,从而可以深入理解无线传感器网络中影响保密率的因素;然后,通过建立一种非合作传感器节点保密率博弈模型,解决传感器节点力图最大化各自保密率的同时,影响整个网络通信的问题;最后,利用演化博弈论中的复制动力学模型,给出传感器节点如何动态地选择各自的发射功率以最大化适应度的演化过程,实现传感器节点保密率自适应调节的机制。

在扩展作者前期工作[453,454]的基础上,本章的工作主要包括以下内容:

(1) 在考虑无线传感器网络存在通信干扰的情况下,通过扩展经典的窃听模型,构建了适用于聚簇无线传感器网络的用于计算传感器节点与其对应的簇头节点之间的保密率模型。

(2) 以传感器节点之间的相互交互为基础,建立了一个非合作保密率博弈模型来反映传感器节点传送数据时的能量消耗。通过最大化传感器节点的各自收益实现能正确选择它们各自的功率策略的目的。

(3) 从演化博弈论角度,得到了能激励传感器节点去寻求具有较高适应度的功率策略,以及揭示哪个功率策略能最终被变异者选择的演化稳定策略,相应地,实现了传感器节点之间保密率的自适应调节。

本章其余章节安排如下:10.2 节介绍相关工作;10.3 节讨论传感器节点之间的干扰模型,并给出如何计算聚簇无线传感器网络中传感器节点与相应的簇头之间保密率的计算公式;10.4 节首先建立传感器节点之间的保密率博弈模型,然后利用演化博弈论中的复制动力学分析传感器节点保密率的变化过程,再给出传感器节点保密率自适应调节算法;10.5 节通过数值实验说明各成本参数对传感器节点保密率博弈模型的影响,也说明了传感器节点保密率自适应调节的过程;10.6 节给出本章小结。

本章涉及的符号含义如下:

ρ 表示无线传感器网络中传感器节点的分布密度。

r_R 表示传感器节点的数据接收范围。

r_I 表示传感器节点的通信干扰范围。

$\sqrt{\sigma}$表示标准方差。

z_I 表示传感器节点的通信干扰区域。

I 表示干扰一个传感器节点通信的其他传感器节点数的最大值。

S 表示包含 M 个传感器节点的集合。

$\mathcal{H}$表示包含 N 个簇头节点的集合。

ε 表示包含 K 个窃听者的集合。

ε_m 表示能窃听传感器节点 S_m 感应数据的窃听者集合。

$G_{H_n}^{S_m}$ 表示传感器节点 S_m 和其对应的簇头节点 H_n 之间的信道增益。

$G_{E_k}^{S_m}$ 表示传感器节点 S_m 和窃听者 $E_k \in \varepsilon_m$ 之间的信道增益。

η^2 表示簇头节点和窃听者之间的热噪声功率。

W 表示每个信道的带宽。

$\hat{S}_i^m$ 表示干扰传感器节点 S_m 信号发送的干扰者,$i \in \{1,2,\cdots,I_m\}$。

$\hat{S}_m$ 表示干扰传感器节点 S_m 信号发送的干扰者的集合。

$C_{H_n}^{S_m}$ 表示传感器节点 S_m 到其对应簇头节点 H_n 的信道容量。

P_m 表示传感器节点 S_m 选择的发射功率。

$\hat{P}_i^m$ 表示干扰节点$\hat{S}_i^m$ 选择的发射功率。

$G_{H_n}^{\hat{S}_i^m}$ 表示干扰节点$\hat{S}_i^m$ 和簇头 H_n 之间的信道增益。

$C_{E_k}^{S_m}$ 表示传感器节点 S_m 到窃听者 $E_k \in \varepsilon_m$ 的信道容量。

$G_{E_k}^{\hat{S}_i^m}$ 表示干扰节点$\hat{S}_i^m$ 和窃听者 $E_k \in \varepsilon_m$ 之间的信道增益。

$C(P_m)$表示选择数据发送功率 P_m 的传感器节点 S_m 和其对应簇头节点 H_n 之间的保密率。

$\mathbb{G}$ 表示本章定义的传感器节点保密率博弈模型。

$\mathcal{P}$表示所有传感器节点可选择功率策略的集合。

$\mathcal{U}$表示传感器节点 S_m选择功率策略 P_m，且它的对手选择功率策略 $P_{\bar{m}}$时的效用集合。

$\mathcal{P}_m$表示传感器节点 S_m 可选择的功率策略 P_m的集合。

$\mu(P_m, P_{\bar{m}})$表示传感器节点 S_m选择功率策略 P_m，且它的对手选择功率策略 $P_{\bar{m}}$ 时的效用。

α 表示用于反映传感器节点在发送数据时消耗能量状况的成本参数。

$\theta_j(t)$表示在时刻 t 选择功率策略 j 的传感器节点在整个无线传感器网络中所占的比例。

$\theta(t)$表示整个无线传感器网络的混合策略。

$\mu_j(t)$表示传感器节点 S_m 在时刻t 选择功率策略 j 的适应度。

$\bar{\mu}(t)$表示整个无线传感器网络在时刻 t 的平均适应度。

$\zeta_j(t)$表示传感器节点 S_m 在时刻t 选择功率策略j 的期望保密率。

$\bar{\zeta}(t)$表示整个无线传感器网络在时刻 t 的平均期望保密率。

$r_j(\theta)$表示选择功率策略 j 的传感器节点的平均策略改变率。

$p_q^j(\theta)$表示传感器节点改变当前功率策略 j 到 q 的概率。

ϕ 表示一个连续可微的概率分布函数。

$\vartheta_j(t)$表示在时刻 t 选择功率策略 j 的传感器节点在缩减种群(Downsized Population)中所占的比例。

$\vartheta(t)$表示在时刻 t 整个缩减种群的状态。

φ_{jq}表示式 $\phi(\mu_j(t)-\mu_q(t))-\phi(\mu_q(t)-\mu_j(t))$的简记符号。

J_{jq}表示 Jacobian 矩阵 $\boldsymbol{J}$ 的元素。

P_{H} 表示高功率策略。

P_{L} 表示低功率策略。

μ_{HH}表示 $\mu(P_{\mathrm{H}}, P_{\mathrm{H}})$的简记符号。

μ_{HL}表示 $\mu(P_{\mathrm{H}}, P_{\mathrm{L}})$的简记符号。

μ_{LH}表示 $\mu(P_{\mathrm{L}}, P_{\mathrm{H}})$的简记符号。

μ_{LL}表示 $\mu(P_{\mathrm{L}}, P_{\mathrm{L}})$的简记符号。

10.2 相关工作

随着各种信道中保密容量研究的深入，当前有很多研究者都在从事如何提高物理层安全的问题。龙航等人[443]较早综述了物理层安全技术的背景和研究现状，说明了目前物理层安全技术的研究以窃听信道容量分析为基础，并对其未来的发展进行了展望，指出了物理层鉴权技术、物理层密钥产生技术和物理层加密技术等研究方向。Mukherjee 等人[455]从 Shannon 和 Wyner 的信息理论安全(Information-theoretic Security)入手，综述了从点对点网络到多天线系统中信息安全策略的演化过程，还介绍了基于物理层的密钥生成协议、信道安全编码方法、基于物理层的消息认证等。李翔宇等人[456]针对中继节点不可信的问题，将中继前后的两个信道等效合并为一个信道后得到联合信道特征，再在联合信道特征的零空间中增加人工噪声，使参与转发的中继节点无法获得有效信息量，从而实现中继物理层安全传输的目的。王亚东等人[457]针对安全编码设计方法对信道条件依赖性强、收发无法共享并具有随机性等问题，提出了一种多天线信道特征投影物理层安全编码算法。卫红权等人[458]基于扰动理论提出了一种适用于频率选择性衰落环境的物理层安全模型，该模型能够通过调节扰动阈值来平衡实际系统的可用性和安全性。陈涛等人[459]通过在次用户的发送信号中加入适当功率的人为噪声，有效地提高了网络的物理层安全性能。罗苗等人[460]针对双向无线协作通信系统的信息论安全问题，提出了一种基于多节点协作波束形成的中继与阻塞混合机制来提高物理层信息传输的安全性。李桥龙和金梁[461]基于无线信道的特征差异，从信息理论安全角度论证了加性随机化权值和乘性随机化权值具有最小信息泄露时应满足的最佳分布，给出了线性随机化预处理模型，设计出了一种实用的物理层安全传输机制。邓浩等人[462]针对传统干扰策略无法有效利用所有协作节点的阻塞功率的问题，提出了一种多节点分组协作干扰以增强无线网络保密率的策略，实现了组内的协作节点能近乎完全利用可用的阻塞总功率的目的。当采用空域加扰实现物理层安全时，构造的多天线加权向量在合法信道上的投影具有恒模特性，窃听者能够利用这一特性截获私密信息，针对这一问题，李明亮等人[463]设计了一种基于空频联合加扰的物理层安全算法。林通等人[464]针对无线多播系统受限于发送方天线数目，整体信道通常不存在零空间，无法利用传统的物理层安全技术保证其安全传输的问题，提出了一种基于多载波的多播系统物理层安全方案。吉江等人[465]提出了基于随机发送参考的多天线系统传输算法，该算法将授权用户的信道分解为多个独立并行的信道，并在其中的一个信道中发送随机化的导频信息，同时对其他信道加密，从而实现物理层信息安全传输的目的。崔波等人[466]针对有限字符输入系统的无线物理层安全传输问题，利用 MIMO 系统的接收天线索引承载信息，通过切换接收天线随机化窃听者的等效信道来保证物理层安全传输。赵耀环等人[467]针对分布式天线的场景，从中间节点中选择一个最佳的节点作为中继，将剩余的其他节点作为协同干扰节点，提出了一种结合最优中继选择和功率分配的物理层安全方案。Wang 等人[468]针对双波传播功率衰退信道(Two-wave with Diffuse Power Fading Channel)，提倡利用最大比合并(Maximal Ratio Combining)技术提高信道的保密率。Hong 和 Chen[469]针对采用合作 MIMO 通信的无线传感器网络，基于信息理论和密码学提出了一种跨层的安全通信模式，为网络中存在的妥协节点攻击(Compromised Nodes Attack)给出了解决方案，从而有效改善了物理层安全。

Hanif 等人[470]针对多用户多天线无线网络，通过线性预编码策略（Linear Precoding Strategies）最大化网络的保密率。Chae 等人[471]利用带保密区域保护的人工噪声技术得到了最优的功率分配，增强了随机无线网络的保密率。

当前，已有研究者利用博弈论研究不同网络中的保密率问题。肖宛阳等人[472]分析了传统物理层安全方法在解决多主体最优化策略求解中存在的问题，给出了基于博弈论的物理层安全模型，分析了基于博弈论的物理层安全建模方法。洪颖等人[473]针对中继自私性导致数据发送中断、源节点保密率降低的问题，提出了一种基于两次报价博弈机制的无线网络安全中继选择方法。都晨辉等人[474]为了使得源节点和协作节点之间取得最优的效益分配，提出了一种基于 Stackelberg 博弈的能效最优报偿及功率分配方案，通过给定优化的协作干扰策略保证了物理层安全。黄开枝等人[475]提出了一种基于演化博弈的物理层安全协作方法，该方法在定义博弈策略（发送人工噪声或信号）和收益（不同策略组合下的安全速率）后，发送端能根据当前网络状态以及协作收益与平均期望收益的差值，不断进行策略调整以最大化收益，再通过求解获得使发送端达到协作稳定策略的条件，使网络从不稳定状态向协作稳定状态演化，从而提高了无线网络的保密率。林胜斌等人[476]针对网络中恶意干扰者通过发送相关干扰破坏合法通信的问题，提出了一种源信号和结构性噪声联合发送的安全传输方法，该方法首先建立发送方和恶意干扰者之间以安全速率为目标函数的连续零和博弈模型，然后根据信道状态确定各自的策略集，并分析策略集对应的纳什均衡，最后利用均衡解指导发送方合理地分配源信号和结构性噪声的功率。吕健体等人[477]针对无线传感器网络中传感器节点之间会产生相互干扰从而影响信道各自的保密率问题，使用非合作博弈建立了无线传感器网络在节点发射功率受限、节点之间存在互相干扰的情况下的博弈模型，得到了传感器节点发射功率的纳什均衡解，以获取最佳的收益。Wang 等人[478]针对包含多个源—目标链路和一个中继节点组成的合作网络，运用拍卖理论有效分配中继节点的能量，提高了网络的保密率。Yuksel 等人[479]为得到优化的源节点保密率，建立了一种源节点和干扰中继之间的非合作博弈模型。Mukherjee 和 Swindlehurst [480]针对 MIMO 窃听信道模型，通过建立以遍历 MIMO 保密率为支付函数的一个零和博弈，得到了可靠且隐蔽的数据发送策略。Han 等人[481]利用友好的干扰者干扰窃听者的通信，提高源节点的保密率，建立了一种 Stackelberg 博弈模型，用以反映源节点和友好干扰者的相互关系。Gabry 等人[482]利用 Stackelberg 博弈分析认知无线电网络中首次发送者（Primary Transmitter）和二次发送者（Secondary Transmitter）之间的合作关系，得到的 Stackelberg 均衡为合作提高认知无线电网络的保密率提供了优化策略。Chu 等人[483]将 Stackelberg 博弈用于解决 MIMO 信道中的功率最小化和保密率最大化问题。Qu 等人[484]运用 Stackelberg 博弈解决 D2D（Device-to-Device）通信网络链路中的功率分配问题，提高了 D2D 通信网络的保密率。Saad[485]和 Fakoorian[486]从合作博弈角度，分别引入联盟博弈和 Kalai-Smorodingsky 讨价还价博弈，研究物理层安全问题。Liu 等人[487]以一个由源节点、目标节点、窃听节点、多个中继节点组成的解码转发网络为研究背景，在考虑分布式中继选择和安全波束成形的基础上分析网络的总保密率，利用联盟博弈对中继选择问题建立的博弈模型降低了计算复杂度。Hou 和 Fu[488]面对无线网络中的合作能有效提高物理层安全的现实情况，利用联盟博弈提出了一种用户有效加入联盟和从联盟中分裂的算法，得到了最大化的物理层保密容量。

与上述相关工作不同的是，本章着重关注的是无线传感器网络中一个传感器节点与其

对应的簇头通信时的保密率，而大部分相关工作主要围绕不同的窃听信道展开研究工作。为此，本章在扩展经典保密率计算公式的基础上得到了能适应聚簇无线传感器网络环境的保密率计算公式。本章将博弈论应用于解决传感器节点力图最大化各自保密率的同时影响整个网络通信的问题，因此，建立了一种非合作传感器节点保密率博弈模型，并基于演化博弈中的复制动力学模型给出了传感器节点如何实现最大化适应度的演化过程，得到了一种实现传感器节点保密率自适应调节的方法。

10.3 系统模型

10.3.1 传感器节点干扰模型

本章考虑静态部署的无线传感器网络，即传感器节点一旦部署完毕将不再移动。其中，每个传感器节点具有唯一的标识，且以节点密度 ρ 部署在二维平面上。每个传感器节点配备具有相同增益的发送和接收天线，并以半双工形式工作，即不允许同时发送和接收通信数据。在相同区域内，传感器节点的接收和干扰范围主要根据源节点和其他相邻节点的发射功率确定，每个传感器节点能根据不同的环境选择不同的发射功率。所有传感器节点的部署满足泊松分布规律。

设 r_R 和 r_I 分别是传感器节点的接收和干扰距离，则 $r_I \approx 2r_R$。通常，一个传感器节点的最大干扰节点数由平均的相邻节点数加上 3 倍的标准差$\sqrt{\sigma}$确定。对任何一个传感器节点，它的干扰区域 z_I 为 πr_I^2，在干扰区域 z_I 中的平均相邻传感器节点数为 ρz_I。在泊松分布中，σ 即为 ρz_I。因此，一个传感器节点的最大干扰节点数为

$$I = \rho z_I + 3\sqrt{\rho z_I} \tag{10-1}$$

10.3.2 聚簇无线传感器网络中的传感器节点保密率

如图 10-1 所示，本章研究的无线传感器网络采用聚簇结构。在该结构中，所有传感器节点被分成不同的簇。每个簇包括 1 个簇头和若干个传感器节点。传感器节点捕获数据后首先将数据传输到同一个簇内的簇头，簇头经数据汇聚后再通过其他簇中的簇头将数据传输到基站。需注意的是，本章研究的是一个传感器节点与其对应的簇头通信时的保密率。

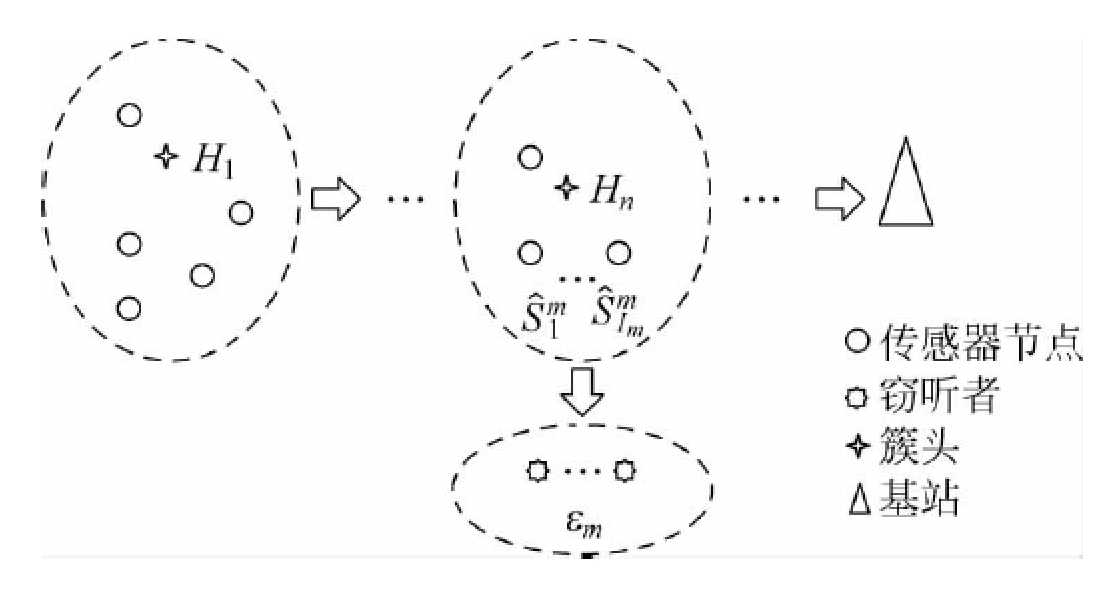


图 10-1 网络模型

记 $S=\{S_1,S_2,\cdots,S_M\}$、$\mathcal{H}=\{H_1,H_2,\cdots,H_N\}$ 和 $\varepsilon=\{E_1,E_2,\cdots,E_K\}$ 分别为包含 M 个传感器节点、N 个簇头和 K 个窃听者的集合。对 $\forall S_m \in S$，存在一个唯一的 $H_n \in \mathcal{H}$，其中 n

值实质由 m 值确定，还存在由多个能窃听传感器节点 S_m 的窃听者 E_k 组成的集合 $\varepsilon_m \subseteq \varepsilon$。记 $G_{H_n}^{S_m}$ 和 $G_{E_k}^{S_m}$ 分别表示从传感器节点 S_m 到簇头 H_n 和窃听者 E_k 的信道增益。为简化模型，假设在簇头端和窃听者端的热噪声功率均为 η^2，每个信道的带宽均为 W。

当多个传感器节点同时发送数据时，将产生信号干扰。对 $\forall S_m \in S$，记 $\hat{S}_m = \{\hat{S}_1^m, \hat{S}_2^m, \cdots, \hat{S}_{I_m}^m\}$ 为包含 I 个干扰 S_m 信号发送的干扰者，其中 I 值由式(10-1)确定。根据经典窃听信道模型计算信道容量的思想可知，从传感器节点 S_m 到对应簇头 H_n 的信道容量为

$$C_{H_n}^{S_m} = W \log_2 \left(1 + \frac{P_m G_{H_n}^{S_m}}{\sum_{i=1}^{I} \hat{P}_i^m G_{H_n}^{\hat{S}_i^m} + \eta^2} \right) \tag{10-2}$$

式中，P_m 为传感器节点 S_m 选择的发射功率；$\hat{P}_i^m$ 为干扰节点 $\hat{S}_i^m$ 选择的发射功率；$G_{H_n}^{\hat{S}_i^m}$ 为干扰节点 $\hat{S}_i^m$ 和簇头 H_n 之间的信道增益。

类似地，从传感器节点 S_m 到窃听者 E_k 的信道容量为

$$C_{E_k}^{S_m} = W \log_2 \left(1 + \frac{P_m G_{E_k}^{S_m}}{\sum_{i=1}^{I} \hat{P}_i^m G_{E_k}^{\hat{S}_i^m} + \eta^2} \right) \tag{10-3}$$

式中，$G_{E_k}^{\hat{S}_i^m}$ 为干扰节点 $\hat{S}_i^m$ 和窃听者 E_k 之间的信道增益。因此传感器节点 S_m 和对应簇头 H_n 之间的保密率为

$$C(P_m) = \left(C_{H_n}^{S_m} - \max_{E_k \in \varepsilon_m} C_{E_k}^{S_m} \right)^+ \tag{10-4}$$

其中，$(x)^+ = \max\{x, 0\}$。

10.4 传感器节点保密率的自适应调节机制

10.4.1 传感器节点保密率博弈模型

定义 10-1 传感器节点保密率博弈模型是一个三元组 $\mathbb{G} = (\mathcal{S}, \mathcal{P}, \mathcal{U})$，其中：

- $\mathcal{S} = \{S_1, S_2, \cdots, S_M\}$ 表示整个无线传感器网络中的传感器节点集合。
- $\mathcal{P} = \prod_{m=1}^{M} \mathcal{P}_m$ 表示所有传感器节点可选择功率策略的集合，$\mathcal{P}_m = \{P_m \mid P_1^m, P_2^m, \cdots, P_L^m\}$ 为传感器节点 S_m 可选择的功率策略 P_m 的集合，L 表示 S_m 可选择的功率策略的个数。
- $\mathcal{U} = \{\mu(P_m, P_{\tilde{m}}) \mid P_m \in \mathcal{P}_m, P_{\tilde{m}} \in \mathcal{P}_{\tilde{m}}, m, \tilde{m} \in \{1, 2, \cdots, M\}\}$ 表示传感器节点 S_m 选择功率策略 P_m 且它的对手选择功率策略 $P_{\tilde{m}}$ 时的效用集合。

根据演化博弈论的观点，可将集合 $\mathcal{S}$ 中的所有传感器节点看作一个种群(Population)，每个传感器节点对应种群中的一个个体。这些个体能根据它们当前的适应度(即期望效用)自动调节各自的策略。也就是说，通过最大化传感器节点的适应度自适应调节各自的保密率。

为反映传感器节点在增大发射功率时提高自身保密率，同时增大自身能量消耗和干扰其他节点的实际情况，定义传感器节点的效用函数为

$$\mu(P_m, P_{\tilde{m}}) = C(P_m) - \alpha P_m \tag{10-5}$$

其中，$C(P_m)$由式(10-4)确定，α 是一个用于反映传感器节点在发送数据时消耗能量状况的成本参数。值得注意的是，选择发射功率 $P_{\tilde{m}}$ 的干扰节点对传感器节点 S_m 的影响已体现在 $C(P_m)$中。

10.4.2 传感器节点保密率的动力学分析

本章利用演化博弈论中的复制动力学分析传感器节点保密率的变化过程。所有传感器节点开始时从各自可用的功率集合中随机选择一种发射功率用于发送通信数据。由于每个传感器节点都希望最大化各自的适应度，所以它们会与无线传感器网络的平均适应度进行比较，调节自身的发射功率，即调节自身的保密率。当整个无线传感器网络的平均适应度高于传感器节点自身的适应度时，它们通过改变自身的发射功率，改变自身的适应度；否则，保持原来的发射功率。记 $\theta_j(t)$为时刻 t 选择功率策略 j 的传感器节点在整个无线传感器网络中所占的比例，则有

$$\sum_{j\in\mathcal{P}_m}\theta_j(t) = 1 \tag{10-6}$$

整个的状态 $\theta(t)$可表示为$[\theta_{P_1^m}(t), \theta_{P_2^m}(t), \cdots, \theta_{P_L^m}(t)]$，可看作整个无线传感器网络的混合策略。记 l 为传感器节点 S_m 的对手 $S_{\tilde{m}}$选择的功率策略。根据参考文献[28]可得，传感器节点 S_m 在时刻 t 选择功率策略 j 的适应度为

$$\mu_j(t) = \sum_{l\in\mathcal{P}_{\tilde{m}}}\theta_l(t)\mu(j, l) \tag{10-7}$$

其中，$\mu(j,l)$由式(10-5)确定。整个无线传感器网络在时刻 t 的平均适应度为

$$\bar{\mu}(t) = \sum_{j\in\mathcal{P}_m}\theta_j(t)\mu_j(t) \tag{10-8}$$

相应地，可定义传感器节点 S_m 在时刻 t 选择功率策略 j 的期望保密率 $\zeta_j(t)$为

$$\zeta_j(t) = \sum_{l\in\mathcal{P}_{\tilde{m}}}\theta_l(t)C(j) \tag{10-9}$$

其中，$C(j)$由式(10-4)确定。整个无线传感器网络在时刻 t 的平均期望保密率 $\bar{\zeta}(t)$为

$$\bar{\zeta}(t) = \sum_{j\in\mathcal{P}_m}\theta_j(t)\zeta_j(t) \tag{10-10}$$

下面分析传感器节点的策略改变率。记 $r_j(\theta)$为选择功率策略 j 的传感器节点的平均策略改变率，$p_q^j(\theta)$为传感器节点改变当前功率策略 j 到 q 的概率，则整个无线传感器网络中改变功率策略 j 到 q 的传感器节点所占比例为 $\theta_j(t)r_j(\theta)p_q^j(\theta)$。因此，从功率策略 j 转换为其他功率策略的传感器节点转出总比例为

$$\sum_{q,q\neq j}\theta_j(t)r_j(\theta)p_q^j(\theta) = \theta_j(t)r_j(\theta)\sum_{q,q\neq j}p_q^j(\theta) = \theta_j(t)r_j(\theta)(1-p_j^j(\theta)) \tag{10-11}$$

从其他功率策略转换到功率策略 j 的传感器节点转入总比例为 $\sum_{q,q\neq j}\theta_q(t)r_q(\theta)p_j^q(\theta)$。

至此，将选择功率策略 j 的传感器节点转入总比例减去转出总比例，可得到整个无线传感器网络中选择功率策略 j 的传感器节点比例变化的动力学方程为

$$\dot{\theta}_j(t) = \sum_{q,q\neq j}\theta_q(t)r_q(\theta)p_j^q(\theta) - \theta_j(t)r_j(\theta)(1-p_j^j(\theta)) = \sum_q\theta_q(t)r_q(\theta)p_j^q(\theta) - \theta_j(t)r_j(\theta) \tag{10-12}$$

式(10-12)中的 $p_j^q(\theta)$决定于选择功率策略 j 和 q 的适应度 $\mu_j(t)$和 $\mu_q(t)$。只有 $\mu_q(t)>\mu_j(t)$时,传感器节点才会从功率策略 j 改变为 q。因此,存在一个连续可微的概率分布函数 $\phi:\mathbb{R}\to[0,1]$,使得

$$p_j^q(\theta)=\begin{cases}\phi(\mu_q(t)-\mu_j(t)), & q\neq j\\ 1-\sum\limits_{q\neq j}\phi(\mu_q(t)-\mu_j(t)), & q=j\end{cases}\tag{10-13}$$

为简化分析过程,令所有传感器节点的平均策略改变率恒等于1,即

$$\forall j\in\mathcal{P}_m,\quad r_j(\theta)\equiv 1\tag{10-14}$$

将式(10-13)和式(10-14)代入式(10-12),则整个无线传感器网络中选择功率策略 j 的传感器节点比例变化的动力学方程为

$$\dot{\theta}_j(t)=\theta_j(t)\sum_{q,q\neq j}\theta_q(t)(\phi(\mu_j(t)-\mu_q(t))-\phi(\mu_q(t)-\mu_j(t)))\tag{10-15}$$

根据参考文献[18]中线性化函数 ϕ 的思想,可设

$$\phi(x)=\beta+\gamma x\tag{10-16}$$

其中,$\beta,\gamma\in\mathbb{R}$ 且 $0\leqslant\beta+\gamma x\leqslant 1$。将式(10-16)代入式(10-15),得到整个无线传感器网络中选择功率策略 j 的传感器节点比例变化的动力学方程为

$$\dot{\theta}_j(t)=2\gamma\theta_j(t)\sum_{q,q\neq j}\theta_q(t)(\mu_j(t)-\mu_q(t))\tag{10-17}$$

式中,γ 为一个影响整个无线传感器网络达到演化稳定策略收敛速度的参数。

10.4.3 传感器节点保密率博弈模型的收敛性和稳定性

引理 10-1　若功率策略 j 是严格占优的,则$\lim\limits_{t\to\infty}\theta_j(t)=1$。

证明　若功率策略 j 是严格占优的,则不管其他传感器节点选择何种功率策略,选择功率策略 j 的传感器节点都能得到比选择其他功率策略更高的适应度。因此,更多的传感器节点会选择功率策略 j,使得选择功率策略 j 的传感器节点比例在整个种群中逐步提高,最终所有的传感器节点均会选择功率策略 j 作为它们的功率策略,即$\lim\limits_{t\to\infty}\theta_j(t)=1$。证毕。

定理 10-1　整个无线传感器网络的混合策略 $\theta(t)$收敛于均衡点。

证明　显然,在无线传感器网络环境中,选择不同的功率策略将得到不同的适应度,这意味着在所有的功率策略中只有一种功率策略具有最高的适应度。将所有的功率策略对应的适应度进行降序排列后,这些降序的适应度可用 $\mu_1(t)>\mu_2(t)>\cdots>\mu_L(t)$ 表示,相应地,整个种群的状态可记为$\lim\limits_{t\to\infty}\theta(t)=[\theta_1(t),\theta_2(t),\cdots,\theta_L(t)]$。由引理10-1,在满足限制条件 $\theta_j(t)\geqslant 0$ 和 $\sum\limits_{j\in\mathcal{P}_m}\theta_j(t)=1$ 前提下,整个无线传感器网络的混合策略最终收敛于

$$\lim_{t\to\infty}\theta(t)=[\theta_1(t),\theta_2(t),\cdots,\theta_L(t)]=[\underbrace{1,0,\cdots,0}_{L}]\tag{10-18}$$

证毕。

定理 10-2　传感器节点保密率博弈模型是演化稳定的。

证明　由于 $\sum\limits_{j\in\mathcal{P}_m}\theta_j(t)=1$,可设

$$\theta_1(t)=1-\theta_2(t)-\cdots-\theta_L(t)\tag{10-19}$$

再将式(10-19)代入式(10-15)后，可得时刻 t 选择功率策略 j 的传感器节点在缩减种群(Downsized Population)中所占的比例的动力学方程为

$$\dot{\vartheta}_j(t) = \vartheta_j(t)\left[\varphi_{jq}(1-\vartheta_j(t)) + \sum_{i=P_2^m, i\neq j}^{P_L^m} \vartheta_i(t)\varphi_{ji}\right] \tag{10-20}$$

其中，$j\in\{P_2^m,\cdots,P_L^m\}$；$\varphi_{jq}=\phi(\mu_j(t)-\mu_q(t))-\phi(\mu_q(t)-\mu_j(t))$，并且

$$\vartheta(t) = [\vartheta_2(t),\vartheta_3(t),\cdots,\vartheta_L(t)] \tag{10-21}$$

表示相应的整个种群状态。由定理 10-2，可得均衡点 $\vartheta^*(t)=[\underbrace{0,0,\cdots,0}_{L-1}]$。这样可以得到 $(L-1)\times(L-1)$ 的 Jacobian 矩阵元素

$$J_{jq} = \left[\frac{\partial\dot{\vartheta}_j(t)}{\partial\vartheta_q(t)}\right]_{\vartheta(t)=\vartheta^*(t)} \tag{10-22}$$

其中，$j,q\in\{P_2^m,\cdots,P_L^m\}$。所以，相应的 Jacobian 矩阵可表示为

$$\boldsymbol{J} = \begin{bmatrix} \varphi_{P_2^m P_1^m} & 0 & \cdots & 0 \\ 0 & \varphi_{P_3^m P_1^m} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \varphi_{P_L^m P_1^m} \end{bmatrix} \tag{10-23}$$

在式(10-23)中，$\varphi_{P_2^m P_1^m},\varphi_{P_3^m P_1^m},\cdots,\varphi_{P_L^m P_1^m}$ 实质上是 Jacobian 矩阵 $\boldsymbol{J}$ 特征值。由引理 10-1，对 $\forall j\in\{P_2^m,\cdots,P_L^m\}$，满足 $\phi(\mu_j(t)-\mu_q(t))=0$，因此可得到

$$\varphi_{jP_1^m} = \phi(\mu_j(t)-\mu_q(t)) - \phi(\mu_q(t)-\mu_j(t)) = -\phi(\mu_q(t)-\mu_j(t)) < 0 \tag{10-24}$$

由文献[489]中的定理 2.7.3 可得均衡点 $\vartheta^*(t)=[\underbrace{0,0,\cdots,0}_{L-1}]$是演化稳定的。证毕。

10.4.4 传感器节点保密率自适应调节算法

在每个传感器节点自适应调节各自保密率的过程中，首先计算各自当前的适应度，然后再与整个无线传感器网络的平均适应度进行比较。若差值大于一个给定的上限值，则根据式(10-17)选择新的功率策略。反复进行该过程，直到整个无线传感器网络达到演化稳定状态，此时每个传感器节点选择的功率策略为演化稳定策略，对应最优的网络效用。在整个演化过程中，传感器节点通过自适应改变选择的功率策略实现保密率的自适应调节。传感器节点保密率自适应调节算法的具体过程如下。

算法 10-1 传感器节点保密率自适应调节算法。

1. 初始化 W、I、η^2 和信道增益等所有参数。
2. $t\leftarrow 0$。
3. 以概率 $\theta(t)=[1/L,1/L,\cdots,1/L]$选择一种功率策略 j。
   //这种“等概率”的选择可以保证传感器节点保密率博弈开始时每个传感器节点具
   //有相同的适应度。
4. 根据式(10-7)和式(10-8)分别计算传感器节点选择功率策略 j 的适应度 $\mu_j(t)$ 和整个无线传感器网络的平均适应度 $\bar{\mu}(t)$。
5. 根据式(10-9)和式(10-10)分别计算传感器节点选择功率策略 j 的期望保密率 $\zeta_j(t)$

和整个无线传感器网络的平均期望保密率 $\overline{\zeta}(t)$。

6. DO WHILE .T.

7. 　IF $|\bar{\mu}(t)-\mu_j(t)|<\tau$　//τ 表示一个给定的下限值

8. 　　EXIT

9. 　ENDIF

10. 　IF $|\bar{\mu}(t)-\mu_j(t)|>\tilde{\tau}$　//$\tilde{\tau}$ 表示一个给定的下限值

11. 　　根据式(10-17)计算功率选择概率 $\theta(t+1)$。

12. 　　以 $\theta(t+1)$ 选择一个新的功率策略 j。

13. 　　根据式(10-7)和式(10-8)分别计算传感器节点选择功率策略 j 的适应度 $\mu_j(t+1)$ 和整个无线传感器网络的平均适应度 $\bar{\mu}(t+1)$。

14. 　　根据式(10-9)和式(10-10)分别计算传感器节点选择功率策略 j 的期望保密率 $\zeta_j(t+1)$ 和整个无线传感器网络的平均期望保密率 $\overline{\zeta}(t+1)$。

15. 　ENDIF

16. 　$t \leftarrow t+1$

17. END DO

18. 返回数组 ζ_j 和 $\overline{\zeta}$。

10.5　实验

由于传感器节点在计算能力、存储能力和能量等方面具有局限性，实验过程中假设每个传感器节点在传感器节点保密率博弈过程中能选择的功率策略集合仅包含两个策略。也就是说，对 $\forall S_m, \mathcal{P}_m=\{P_H, P_L\}$，其中 P_H 和 P_L 分别表示高功率策略和低功率策略。

传感器节点保密率博弈模型中的参数根据 IEEE 802.15.4 物理层规范进行设置。其中，$\rho=0.01$，$W=2\text{MHz}$，$P_H=30\text{mW}$，$P_L=10\text{mW}$，$\sigma^2=-112\text{dBm}$，$G_{H_n}^{S_m}=1$，$G_{E_k}^{S_m}=0.6$。由于传感器节点选择的功率越大，产生的干扰范围越大，所以分别设置 $r_H=50\text{m}$ 和 $r_L=10\text{m}$，其中，r_H 和 r_L 分别表示传感器节点选择功率策略 P_H 和 P_L 产生的干扰半径。另外，根据经验值假设干扰者的工作概率为 0.01。这样结合式(10-1)，可得到选择功率策略 P_H 和 P_L 的干扰节点数量分别为 $0.01\times(\rho\pi r_H^2+3\sqrt{\rho\pi r_H^2})$ 和 $0.01\times(\rho\pi r_L^2+3\sqrt{\rho\pi r_L^2})$。

为方便描述，记 $\theta_H(t)$ 为整个无线传感器网络中传感器节点在时刻 t 选择功率策略 P_H 所占的比例，则整个无线传感器网络中传感器节点在时刻 t 选择功率策略 P_L 所占的比例为 $1-\theta_H(t)$。由式(10-7)，传感器节点在时刻 t 选择功率策略 P_H 和 P_L 的适应度 $\mu_H(t)$ 和 $\mu_L(t)$ 分别为

$$\mu_H(t)=\theta_H(t)\mu(P_H,P_H)+(1-\theta_H(t))\mu(P_H,P_L) \tag{10-25}$$

$$\mu_L(t)=\theta_H(t)\mu(P_L,P_H)+(1-\theta_H(t))\mu(P_L,P_L) \tag{10-26}$$

由式(10-8)，整个无线传感器网络在时刻 t 的平均适应度为

$$\bar{\mu}(t)=\theta_H(t)\mu_H(t)+(1-\theta_H(t))\mu_L(t) \tag{10-27}$$

由式(10-9)，传感器节点在时刻 t 选择功率策略 P_H 和 P_L 的期望保密率 $\zeta_H(t)$ 和 $\zeta_L(t)$ 分别为

$$\zeta_H(t)=\theta_H(t)\zeta_{HH}+(1-\theta_H(t))\zeta_{HL} \tag{10-28}$$

$$\zeta_{L}(t)=\theta_{H}(t)\zeta_{LH}+(1-\theta_{H}(t))\zeta_{LL} \tag{10-29}$$

其中 ζ_{uv}，$u,v\in\{H,L\}$，表示当干扰节点选择功率策略 P_v 时，一个选择功率策略 P_u 的传感器节点的保密率，其值由式(10-4)计算得到。由式(10-10)，整个无线传感器网络在时刻 t 的平均期望保密率为

$$\overline{\zeta}(t)=\theta_{H}(t)\zeta_{H}(t)+(1-\theta_{H}(t))\zeta_{L}(t) \tag{10-30}$$

由算法 10-1，图 10-2 和图 10-3 给出了当成本参数 α 值为 3 时传感器节点的自适应调节过程。在图 10-2 中，P_H 是传感器节点保密率博弈模型的演化稳定策略。从中可以看出，即使最初选择功率策略 P_H 的比例只有 0.5%，当 t 值大于 40 后，$\theta_H(t)$ 趋向于稳定值 1。这意味着传感器节点选择功率策略 P_H 的适应度总是高于选择 P_L 的适应度，因此所有的传感器节点经过自适应调节最终均选择 P_H 作为自己的功率策略。此时，如图 10-3 所示，$\zeta_H(t)$ 收敛后的极限值约为 132.9976，$\zeta_L(t)$ 收敛后的极限值约为 48.4474，$\overline{\zeta}(t)$ 收敛后的极限值约为 132.9148。

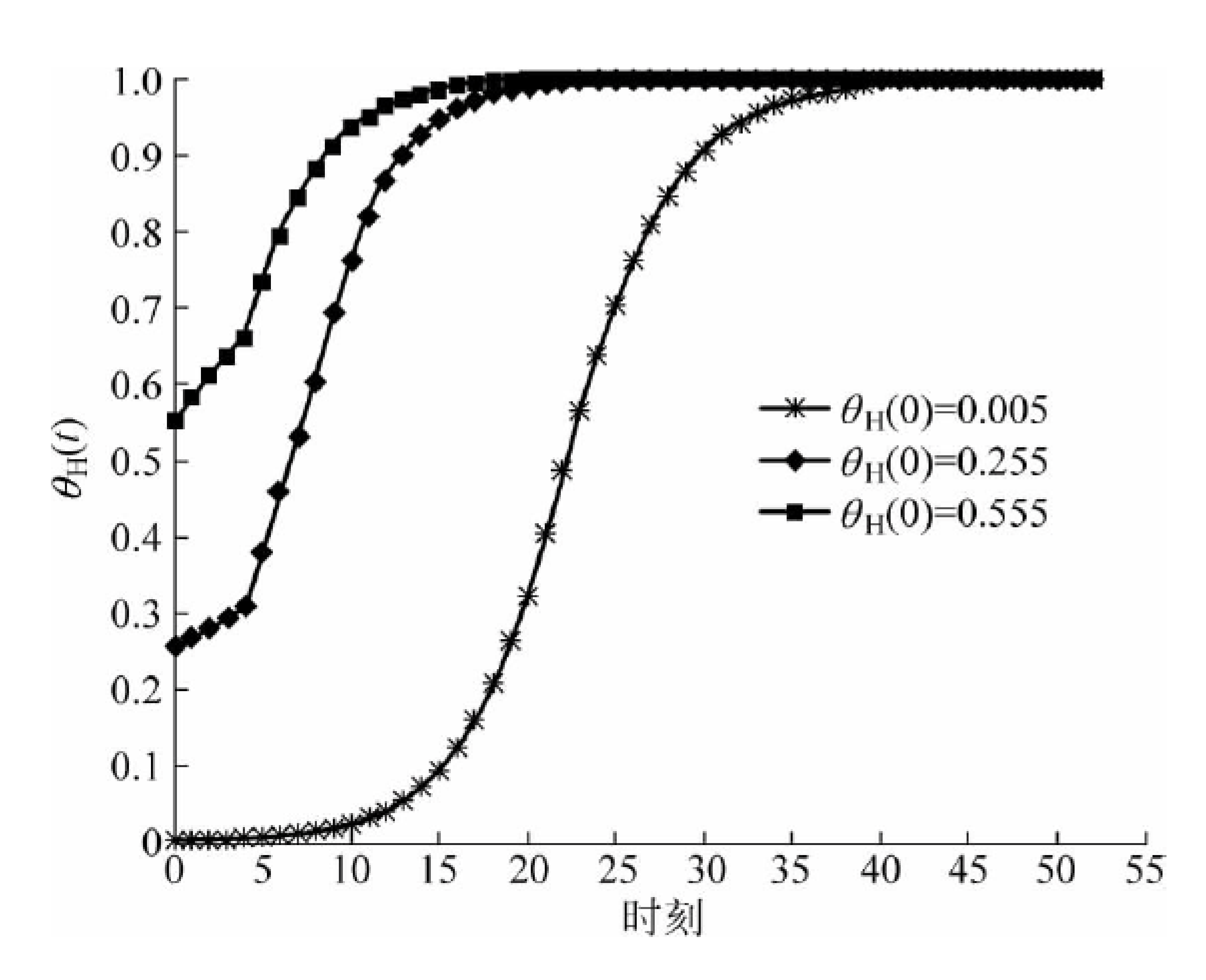


图 10-2 $\alpha=3$ 时的 $\theta_H(t)$ 演化趋势

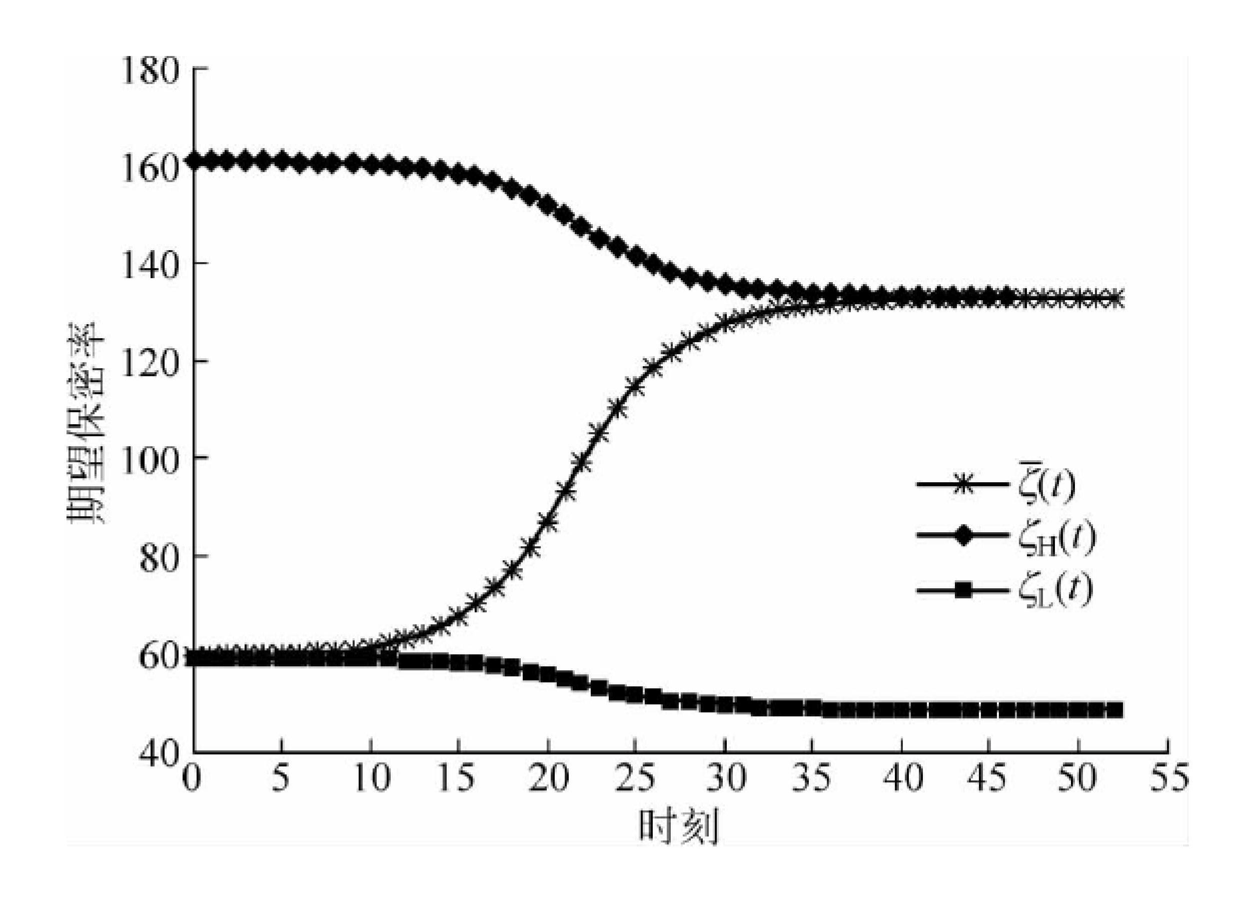


图 10-3 $\alpha=3$ 时的期望保密率演化趋势

由算法1,图10-4和图10-5给出了当成本参数α值为4.5时传感器节点的自适应调节过程。在图10-4中,传感器节点保密率博弈模型存在一个演化稳定混合策略(约为[0.6841,0.3158])。达到该均衡点意味着约68.41%的传感器节点选择P_H,而约31.58%的传感器节点选择P_L。此时,传感器节点选择P_H或P_L具有相同的适应度,但从图10-5可看出,$\zeta_H(t)$收敛后的极限值约为141.8309,$\zeta_L(t)$收敛后的极限值约为51.8309,$\bar{\zeta}(t)$收敛后的极限值约为113.3949。这说明虽然选择功率策略P_H的期望保密率值要高,但并非所有的传感器节点都选择P_H作为自身的功率策略。

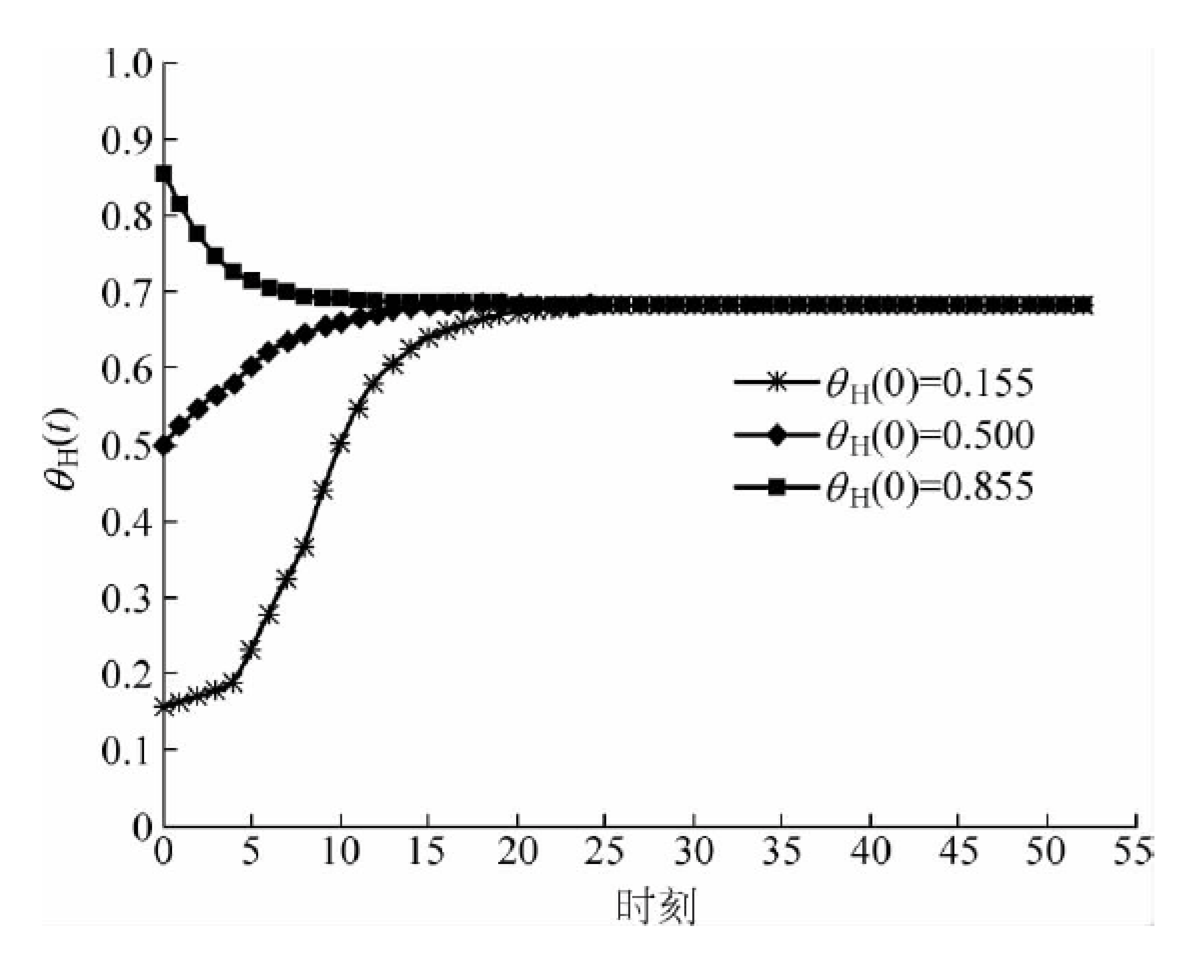


图10-4 α=4.5时的$\theta_H(t)$演化趋势

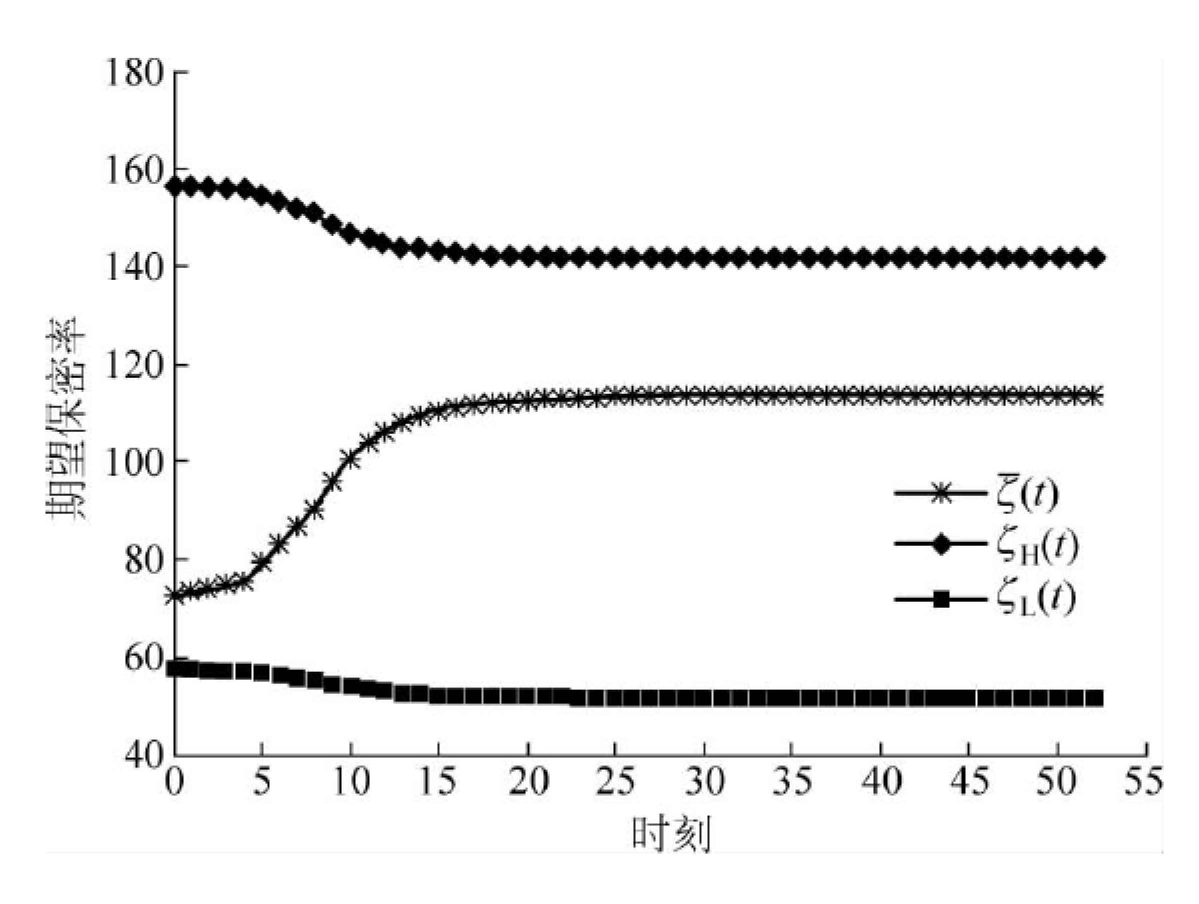


图10-5 α=4.5时的期望保密率演化趋势

由算法1,图10-6和图10-7给出了当成本参数α值为6时传感器节点的自适应调节过程。在图10-6中,P_L是传感器节点保密率博弈模型的演化稳定策略。从中可以看出,即使最初选择功率策略P_H的比例达到99.5%,当t值大于30后,$\theta_H(t)$趋向于稳定值0。这意味着传感器节点选择P_L的适应度总是高于选择P_H的适应度,因此,所有的传感器节点经过自适应调节最终均选择P_L作为自己的功率策略。此时,如图10-7所示,$\zeta_H(t)$收敛后的极限值约为161.0138,$\zeta_L(t)$收敛后的极限值约为59.1789,$\bar{\zeta}(t)$收敛后的极限值约为

59.1994。这说明虽然选择功率策略 P_H 的期望保密率值要高，但所有的传感器节点仍选择 P_L 作为自身的功率策略。

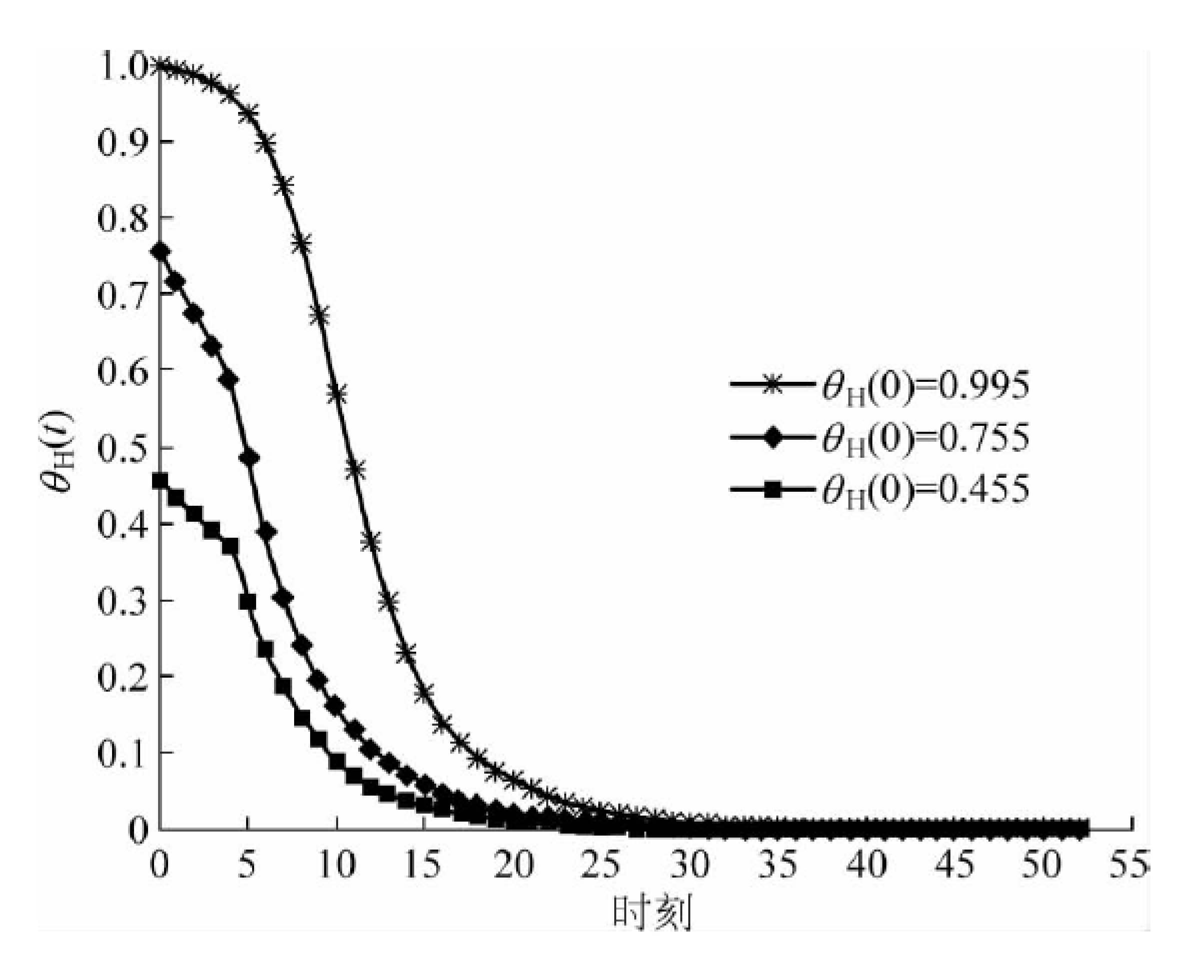


图 10-6 $\alpha=6$ 时的 $\theta_H(t)$ 演化趋势

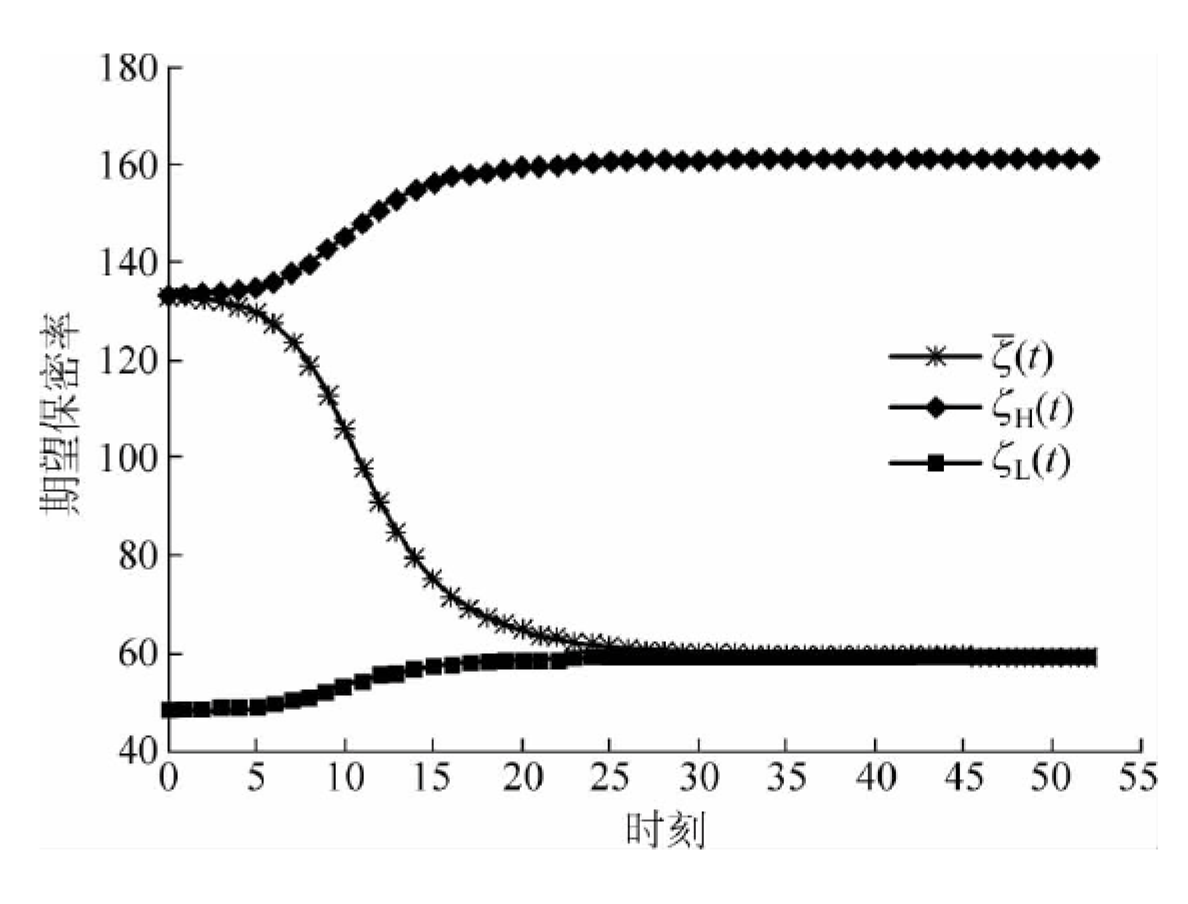


图 10-7 $\alpha=6$ 时的期望保密率演化趋势

10.6 小结

本章基于经典窃听信道模型扩展的保密率公式可适应聚簇无线传感器网络环境，构建的传感器节点保密率博弈模型能反映传感器节点的交互过程。基于演化博弈论中复制动力学原理的传感器节点比例变化动力学方程，能使传感器节点根据自身当前的适应度和无线传感器网络的平均适应度，动态选择自身的功率策略，是一种自适应调节传感器节点保密率的方法。实验进一步阐明了该方法的机理，为利用物理层安全技术保证无线传感器网络数据的保密性提供了新途径。

参考文献

[1] 孙利民,李建中,陈渝,等. 无线传感器网络[M]. 北京:清华大学出版社, 2005.

[2] Akyildiz I F, Su W, Sankarasubramaniam Y, *et al*. Wireless sensor networks: A survey[J]. Computer Networks, 2002, 38(4): 393-422.

[3] 任丰原,黄海宁,林闯. 无线传感器网络[J]. 软件学报, 2003, (7): 1282-1291.

[4] 崔莉,鞠海玲,苗勇,等. 无线传感器网络研究进展[J]. 计算机研究与发展, 2005, (1): 163-174.

[5] Yick J, Mukherjee B, Ghosal D. Wireless sensor network survey[J]. Computer Networks, 2008, 52(12): 2292-2330.

[6] 唐宏,谢静,鲁玉芳,等. 无线传感器网络原理及应用[M]. 北京:人民邮电出版社, 2010.

[7] 王汝传,孙力娟,郭剑,等. 无线传感器网络技术及其应用[M]. 北京:人民邮电出版社, 2011.

[8] 余成波,李洪兵,陶红艳. 无线传感器网络实用教程[M]. 北京:清华大学出版社, 2012.

[9] 姚向华,杨新宇,易劲刚,等. 无线传感器网络原理与应用[M]. 北京:高等教育出版社, 2012.

[10] Baronti P, Pillai P, Chook V, *et al*. Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards[J]. Computer Communications, 2007, 30(7): 1655-1695.

[11] Chen X, Makki K, Yen K, *et al*. Sensor network security: A survey[J]. IEEE Communications Surveys & Tutorials, 2009, 11(2): 52-73.

[12] 裴庆祺,沈玉龙,马建峰. 无线传感器网络安全技术综述[J]. 通信学报, 2007, 28(8): 113-122.

[13] Zhou Y, Fang Y, Zhang Y. Securing wireless sensor networks: A survey[J]. IEEE Communications Surveys & Tutorials, 2008, 10(3): 6-28.

[14] 王潮,胡广跃,张焕国. 无线传感器网络的轻量级安全体系研究[J]. 通信学报, 2012, (2): 30-35.

[15] Fudenberg D, Tirole J. Game Theory[M]. London: The MIT Press, 1991.

[16] 李光久. 博弈论基础教程[M]. 北京:化学工业出版社, 2005.

[17] 肖条军. 博弈论及其应用[M]. 上海:上海三联书店, 2004.

[18] Weibull J W. Evolutionary Game Theory[M]. Cambridge: MIT Press, 1995.

[19] 范如国. 博弈论[M]. 武汉:武汉大学出版社, 2011.

[20] Başar T, Olsder G J. Dynamic Noncooperative Game Theory[M]. New York: Academic Press, 1999.

[21] Peters H. Game Theory-A Multi-Leveled Approach[M]. Berlin: Springer-Verlag, 2008.

[22] Berr F. Stackelberg equilibria in managerial delegation games[J]. European Journal of Operational Research, 2011, 212(2): 251-262.

[23] Boche H, Schubert M. A generalization of Nash bargaining and proportional fairness to log-convex utility sets with power constraints[J]. IEEE Transactions on Information Theory, 2011, 57(6): 3390-3404.

[24] 谢识予. 经济博弈论[M]. 上海:复旦大学出版社, 2007.

[25] Watanabe T, Yamato T. A choice of auction format in seller cheating: a signaling game analysis[J]. Economic Theory, 2008, 36(1): 57-80.

[26] Zhuang J, Bier V M, Alagoz O. Modeling secrecy and deception in a multiple-period attacker-defender signaling game[J]. European Journal of Operational Research, 2010, 203(2): 409-418.

[27] Sergiu H, Andreu M. Cooperation: game-theoretic approaches[M]. Berlin: Springer, 1997.

[28] Fudenberg D, Levine D K. The Theory of Learning in Games [M]. Cambridge: MIT Press, 1998.

[29] Isaacs R. Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization[M]. New York: John Wiley and Sons, 1965.

[30] Friedman A. On the definition of differential games and the existence of value and saddle points[J].

Journal of Differential Equations，1970，7(1)：69-91.
[31] Friedman A. Existence of value and of saddle point for differential games of survival[J]. Journal of Differential Equations，1970，7(1)：111-125.
[32] Friedman A. Existence of value and of saddle point for differential games of pursuit and evasion[J]. Journal of Differential Equations，1970，7(1)：92-110.
[33] 张嗣瀛. 微分对策[M]. 北京：科学出版社，1987.
[34] 李登峰. 微分对策及其应用[M]. 北京：国防工业出版社，2000.
[35] 黄涛. 博弈论教程——理论・应用[M]. 北京：首都经济贸易大学出版社，2004.
[36] Bressan A. Noncooperative differential games[J]. Milan Journal of Mathematics，2011，79(2)：357-427.
[37] De Marco G，Romaniello M. A dynamic game of coalition formation under ambiguity[J]. Soft Computing，2011，15(4)：637-644.
[38] Gharehshiran O N，Krishnamurthy V. Coalition formation for bearings-only localization in sensor networks-A cooperative game approach[J]. IEEE Transactions on Signal Processing，2010，58(8)：4322-4338.
[39] Li C L，Yang Z，Li J，*et al*. A relaying incentive scheme for multihop cellular networks based on coalition game with externalities[J]. Wireless Personal Communications，2011，58(4)：785-805.
[40] Liang X，Xiao Y. Studying bio-inspired coalition formation of robots for detecting intrusions using game theory[J]. IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics，2010，40(3)：683-693.
[41] Fadlullah Z M，Taleb T，Vasilakos A V，*et al*. DTRAB：Combating against attacks on encrypted protocols through traffic-feature analysis[J]. IEEE/ACM Transactions on Networking，2010，18(4)：1234-1247.
[42] Wang Y，Nakao A，Vasilakos A V，*et al*. P2P soft security：On evolutionary dynamics of P2P incentive mechanism[J]. Computer Communications，2011，34(3)：241-249.
[43] Agah A，Basu K，Das S K. Preventing DoS attack in sensor networks：A game theoretic approach[C]. Proc. of IEEE International Conference on Communications，2005：3218-3222.
[44] Agah A，Das S K，Basu K. A game theory based approach for security in wireless sensor networks[C]. Proc. of IEEE International Performance，Computing and Communications Conference，2004：259-263.
[45] Agah A，Das S K. Preventing DoS attacks in wireless sensor networks：A repeated game theory approach[J]. International Journal of Network Security，2007，5(2)：145-153.
[46] Yang L，Mu D，Cai X. Preventing dropping packets attack in sensor networks：A game theory approach[J]. Wuhan University Journal of Natural Sciences，2008，13(5)：631-635.
[47] Mohi M，Movaghar A，Zadeh P M. A bayesian game approach for preventing DoS attacks in wireless sensor networks[C]. Proc. of WRI International Conference on Communications and Mobile Computing，2009：507-511.
[48] Mccune J M，Shi E，Perrig A，*et al*. Detection of denial-of-message attacks on sensor network broadcasts[C]. Proc. of IEEE Symposium on Security and Privacy，2005：64-78.
[49] Reddy Y B. A game theory approach to detect malicious nodes in wireless sensor networks[C]. Proc. of Third International Conference on Sensor Technologies and Applications，2009：462-468.
[50] Huang J，Liao I，Chung Y，*et al*. Shielding wireless sensor network using Markovian intrusion detection system with attack pattern mining[J]. Information Sciences，2013，231：32-44.
[51] Chen L，Leneutre J. Fight jamming with jamming-A game theoretic analysis of jamming attack in wireless networks and defense strategy[J]. Computer Networks，2011，55(9)：2259-2270.

[52] 李奕男,钱志鸿,刘影,等. 基于博弈论的移动 Ad hoc 网络入侵检测模型[J]. 电子与信息学报, 2010, (9): 2245-2248.
[53] Kantzavelou I, Katsikas S. A game-based intrusion detection mechanism to confront internal attackers[J]. Computers & Security, 2010, 29(8): 859-874.
[54] Shen S, Li Y, Xu H, *et al*. Signaling game based strategy of intrusion detection in wireless sensor networks[J]. Computers & Mathematics with Applications, 2011, 62(6): 2404-2416.
[55] 曹晓梅,韩志杰,陈贵海. 基于流量预测的传感器网络拒绝服务攻击检测方案[J]. 计算机学报, 2007, (10): 1798-1805.
[56] Min W, Keecheon K. Intrusion detection scheme using traffic prediction for wireless industrial networks[J]. Journal of Communications and Networks, 2012, 14(3): 310-318.
[57] 周集良,李彩霞,曹奇英. 基于 WSNs 安全协议的入侵检测系统研究[J]. 计算机应用研究, 2009, (11): 4319-4321.
[58] 王骐,王殊,孟中楼. 分布式入侵检测系统的融合算法[J]. 华中科技大学学报(自然科学版), 2009, (9): 49-52.
[59] 刘宁,范训礼,赵建华. 一种无线传感器网络入侵检测系统模型[J]. 西南科技大学学报, 2009, (1): 78-81.
[60] 肖政宏,陈志刚,李庆华. WSN 中基于分布式机器学习的异常检测仿真研究[J]. 系统仿真学报, 2011, (1): 181-187.
[61] 王颖,李国瑞. 基于分组的无线传感器网络入侵检测方案[J]. 传感技术学报, 2009, (6): 878-882.
[62] 王骐,王殊,孟中楼. 无线传感器网络中一种基于接收功率异常的入侵检测算法[J]. 计算机科学, 2009, (3): 34-37.
[63] 张红莉,黄守明. 一种基于 MA 的无线传感器网络 IDS 模型研究[J]. 计算机工程与科学, 2010, (5): 18-20.
[64] 韩志杰,张玮玮,陈志国. 基于 Markov 的无线传感器网络入侵检测机制[J]. 计算机工程与科学, 2010, (9): 27-29.
[65] 祝琦,宋如顺,姚永仙. 无线传感器网络中基于 SVM 的合作型入侵检测系统[J]. 计算机应用研究, 2010, (4): 1489-1492.
[66] 汪淑丽. 基于支持向量机的无线传感器网络的入侵检测系统[J]. 传感器与微系统, 2012, (7): 73-76.
[67] Rajasegarar S, Leckie C, Bezdek J C, *et al*. Centered hyperspherical and hyperellipsoidal one-class support vector machines for anomaly detection in sensor networks[J]. IEEE Transactions on Information Forensics and Security, 2010, 5(3): 518-533.
[68] Wang S S, Yan K Q, Wang S C, *et al*. An integrated intrusion detection system for cluster-based wireless sensor networks[J]. Expert Systems with Applications, 2011, 38(12): 15234-15243.
[69] 王新胜,詹永照,王良民. 无线传感器网络选择性传递攻击协同防御方法[J]. 江苏大学学报(自然科学版), 2011, (4): 449-455.
[70] 易晓梅,吴鹏,刘丽娟,等. 基于 PSO-RBF 无线传感器网络入侵检测技术研究[J]. 传感器与微系统, 2011, (9): 9-11.
[71] 刘宁,赵建华. 基于生物免疫的无线传感器网络入侵检测系统[J]. 桂林电子科技大学学报, 2011, (2): 138-141.
[72] 刘宁,赵建华. 应用免疫原理的无线传感器网络入侵检测系统[J]. 计算机工程与应用, 2011, (15): 80-82.
[73] 陈珊珊,杨庚,陈生寿. 基于 LEACH 协议的 Sybil 攻击入侵检测机制[J]. 通信学报, 2011, (8): 143-149.
[74] Salmon H M, de Farias C M, Loureiro P, *et al*. Intrusion Detection System for Wireless Sensor

Networks Using Danger Theory Immune-Inspired Techniques[J]. International Journal of Wireless Information Networks，2013，20(1)：39-66.

[75] Choi S，Eom H，Jung E. Securing wireless sensor networks against broadcast service attacks[J]. International Journal of Computers and Applications，2012，34(3)：185-191.

[76] 余小华，黄灿辉，陈瑛. 一种蚁群优化的 WSN 分布式入侵检测模型[J]. 计算机工程与应用，2012，(9)：78-82.

[77] 胡志鹏，魏立线，申军伟，等. 基于核 Fisher 判别分析的无线传感器网络入侵检测算法[J]. 传感技术学报，2012，(2)：246-250.

[78] 范荣真. 基于局部联系的无线传感网络异常入侵检测[J]. 微电子学与计算机，2012，(3)：113-116.

[79] Ho J W，Wright M，Das S K. Zone trust：Fast zone-based node compromise detection and revocation in wireless sensor networks using sequential hypothesis testing[J]. IEEE Transactions on Dependable and Secure Computing，2012，9(4)：494-510.

[80] 毛郁欣. 基于计数器对称加密的无线传感器网络入侵检测算法[J]. 通信学报，2011，32(9A)：211-219.

[81] 石进，陆音，谢立. 基于博弈理论的动态入侵响应[J]. 计算机研究与发展，2008，(5)：747-757.

[82] Chen L，Leneutre J. A game theoretical framework on intrusion detection in heterogeneous networks [J]. IEEE Transactions on Information Forensics and Security，2009，4(2)：165-178.

[83] Dong R，Liu L，Liu J，*et al*. Intrusion detection system based on payoff matrix for wireless sensor networks[C]. Proc. of 3rd International Conference on Genetic and Evolutionary Computing，2009：3-6.

[84] 周四清，李志艳，刘田. 无线传感器网络入侵检测的重复博弈建模研究[J]. 计算机工程与应用，2009，(3)：119-123.

[85] 严辉，沈士根，曹奇英. Ad Hoc 网络中基于重复博弈的攻击预测模型[J]. 计算机工程，2012，(7)：110-112.

[86] 陈行，陶军. 无线网络中基于贝叶斯博弈模型的入侵检测算法研究[J]. 通信学报，2010，31(2)：107-112，119.

[87] Rafsanjani M K，Aliahmadipour L，Javidi M M. A hybrid intrusion detection by game theory approaches in MANET[J]. Indian Journal of Science and Technology，2012，5(2)：2123-2131.

[88] 曹晖，王青青，马义忠，等. 基于静态贝叶斯博弈的攻击预测模型[J]. 计算机应用研究，2007，(10)：122-124.

[89] 曹晖，王青青，马义忠，等. 基于动态贝叶斯博弈的攻击预测模型[J]. 计算机应用，2007，(6)：1545-1547.

[90] 王静，袁凌云，夏幼明，等. 基于激励机制的贝叶斯博弈防御模型[J]. 微型机与应用，2011，(10)：66-68.

[91] Zhu J，Liu Y，Yang X，*et al*. Dynamic game based intrusion response model[J]. Journal of Computational Information Systems，2010，6(7)：2199-2211.

[92] Liang X，Xiao Y. Game theory for network security[J]. IEEE Communications Surveys and Tutorials，2013，15(1)：472-486.

[93] Manshaei M H，Zhu Q，Alpcan T，*et al*. Game theory meets network security and privacy[J]. ACM Computing Surveys，2013，Vol. 45，Article 25，39 pages.

[94] Shen S，Yue G，Cao Q，*et al*. A survey of game theory in Wireless Sensor Networks security[J]. Journal of Networks，2011，6(3)：521-532.

[95] Javidi M M，Aliahmadipour L. Game theory approaches for improving intrusion detection in MANETs[J]. Scientific Research and Essays，2011，6(31)：6535-6539.

[96] 赵柳榕，梅姝娥，仲伟俊. 虚拟专用网和入侵检测系统最优配置策略的博弈分析[J]. 管理工程学报，

2014，(4)：187-192.
[97] Bedi H，Shiva S，Roy S. A game inspired defense mechanism against distributed denial of service attacks[J]. Security and Communication Networks，2014，7(12)：2389-2404.
[98] Shamshirband S，Patel A，Anuar N B，*et al*. Cooperative game theoretic approach using fuzzy Q-learning for detecting and preventing intrusions in wireless sensor networks[J]. Engineering Applications of Artificial Intelligence，2014，32：228-241.
[99] Moosavi H，Bui F M. A game-theoretic framework for robust optimal intrusion detection in wireless sensor networks[J]. IEEE Transactions on Information Forensics and Security，2014，9(9)：1367-1379.
[100] Zonouz S A，Khurana H，Sanders W H，*et al*. RRE：A game-theoretic intrusion response and recovery engine[J]. IEEE Transactions on Parallel and Distributed Systems，2014，25(2)：395-406.
[101] 刘玉枚，杨寿保，路卫娜. P2P环境中基于信号博弈论的资源定价机制[J]. 华中科技大学学报(自然科学版)，2007，(S2)：40-43.
[102] 陈亚睿，田立勤，杨扬. 云计算环境下基于动态博弈论的用户行为模型与分析[J]. 电子学报，2011，(8)：1818-1823.
[103] Patcha A，Park J. A game theoretic formulation for intrusion detection in mobile Ad Hoc networks[J]. International Journal of Network Security，2006，2(2)：131-137.
[104] Wang W，Chatterjee M，Kwiat K. Coexistence with malicious nodes：A game theoretic approach[C]. Proc. of International Conference on Game Theory for Networks，2009：277-286.
[105] Estiri M，Khademzadeh A. A theoretical signaling game model for intrusion detection in wireless sensor networks[C]. Proc. of 14th International Telecommunications Network Strategy and Planning Symposium，2010：1-6.
[106] Li F，Yang Y，Wu J. Attack and flee：Game-theory-based analysis on interactions among nodes in MANETs[J]. IEEE Transactions on Systems，Man，and Cybernetics，Part B：Cybernetics，2010，40(3)：612-622.
[107] Maia S L F，Silva E R，Guardieiro P R. A new optimization strategy proposal for multi-copy forwarding in energy constrained DTNs[J]. IEEE Communications Letters，2014，18(9)：1623-1626.
[108] Paramasivan B，Prakash M J V，Kaliappan M. Development of a secure routing protocol using game theory model in mobile ad hoc networks[J]. Journal of Communications and Networks，2015，17(1)：75-83.
[109] Farooqi A H，Khan F A，Wang J，*et al*. A novel intrusion detection framework for wireless sensor networks[J]. Personal and Ubiquitous Computing，2013，17(5)：907-919.
[110] Younis O，Krunz M，Ramasubramanian S. Node clustering in wireless sensor networks：Recent developments and deployment challenges[J]. IEEE Network，2006，20(3)：20-25.
[111] Yu H，Shen Z，Miao C，*et al*. A survey of trust and reputation management systems in wireless communications[J]. Proceedings of the IEEE，2010，98(10)：1755-1772.
[112] Alpcan T，Başar T. Network Security：A Decision and Game-Theoretic Approach[M]. Cambridge：Cambridge University Press，2010.
[113] Zeng Y，Xiang K，Li D，*et al*. Directional routing and scheduling for green vehicular delay tolerant networks[J]. Wireless Networks，2013，19(2)：161-173.
[114] Shin H M，Lee C Y. Optimal rate allocation and QoS-sensitive admission control in wireless integrated networks[J]. Wireless Networks，2011，17(1)：231-246.
[115] 沈士根，马绚，蒋华，等. 基于演化博弈论的WSNs信任决策模型与动力学分析[J]. 控制与决策，

2012，27(8)：1133-1138.
[116] 荆琦，唐礼勇，陈钟. 无线传感器网络中的信任管理[J]. 软件学报，2008，(7)：1716-1730.
[117] Momani M. Trust models in wireless sensor networks：A survey[C]. Proc. of Communications in Computer and Information Science，2010：37-46.
[118] Esch J. A survey of trust and reputation management systems in wireless communications[J]. Proceedings of the IEEE，2010，98(10)：1752-1754.
[119] Lopez J，Roman R，Agudo I，*et al*. Trust management systems for wireless sensor networks：Best practices[J]. Computer Communications，2010，33(9)：1086-1093.
[120] Yu Y，Li K，Zhou W，*et al*. Trust mechanisms in wireless sensor networks：Attack analysis and countermeasures[J]. Journal of Network and Computer Applications，2012，35(3)：867-880.
[121] 吕林涛，洪磊，张娜. 面向无线传感器网络的分层路由信任模型[J]. 计算机工程，2010，(23)：101-103.
[122] Maarouf I，Baroudi U，Naseer A R. Efficient monitoring approach for reputation system-based trust-aware routing in wireless sensor networks[J]. IET Communications，2009，3(5)：846-858.
[123] Leligou H C，Trakadas P，Maniatis S，*et al*. Combining trust with location information for routing in wireless sensor networks[J]. Wireless Communications and Mobile Computing，2012，12(12)：1091-1103.
[124] 莫英红，钟诚，唐金辉，等. 基于功能信任的无线传感器网络安全数据融合方法[J]. 小型微型计算机系统，2011，(1)：80-84.
[125] 王建萍，李明，周贤伟. 基于声誉和信任组的无线传感器网络实体认证研究[J]. 传感技术学报，2008，(10)：1780-1784.
[126] Boukerch A，Xu L，El-Khatib K. Trust-based security for wireless ad hoc and sensor networks[J]. Computer Communications，2007，30(11-12)：2413-2427.
[127] 黄廷磊，李小龙. 传感器网络中一种信任管理机制[J]. 桂林电子科技大学学报，2010，(5)：428-431.
[128] 董慧慧，郭亚军. 一种基于节点多角度信任的无线传感器网络[J]. 计算机科学，2009，(9)：43-45.
[129] He D，Chen C，Chan S，*et al*. ReTrust：Attack-resistant and lightweight trust management for medical sensor networks[J]. IEEE Transactions on Information Technology in Biomedicine，2012，16(4)：623-632.
[130] Bao F，Chen I，Chang M，*et al*. Hierarchical trust management for wireless sensor networks and its applications to trust-based routing and intrusion detection[J]. IEEE Transactions on Network and Service Management，2012，9(2)：169-183.
[131] 张乐君，邓鑫，国林，等. 基于关联度分析的 WSN 节点信任模型研究[J]. 电子科技大学学报，2015，(1)：106-111.
[132] Aivaloglou E，Gritzalis S. Hybrid trust and reputation management for sensor networks[J]. Wireless Networks，2010，16(5)：1493-1510.
[133] Zhan G，Shi W，Deng J. SensorTrust：A resilient trust model for wireless sensing systems[J]. Pervasive and Mobile Computing，2011，7(4)：509-522.
[134] Mármol F G，Pérez G M. Providing trust in wireless sensor networks using a bio-inspired technique [J]. Telecommunication Systems，2011，46(2)：163-180.
[135] Jiang J，Han G，Wang F，*et al*. An efficient distributed trust model for wireless sensor networks [J]. IEEE Transactions on Parallel and Distributed Systems，2015，26(5)：1228-1237.
[136] Ren Y，Zadorozhny V I，Oleshchuk V A，*et al*. A novel approach to trust management in unattended wireless sensor networks[J]. IEEE Transactions on Mobile Computing，2014，13(7)：1409-1423.

[137] Chae Y, Dipippo L C, Sun Y L. Trust management for defending on-off attacks[J]. IEEE Transactions on Parallel and Distributed Systems, 2015, 26(4): 1178-1191.
[138] Zhou P, Jiang S, Irissappane A, *et al*. Toward energy-efficient trust system through watchdog optimization for WSNs[J]. IEEE Transactions on Information Forensics and Security, 2015, 10(3): 613-625.
[139] Zhu C, Nicanfar H, Leung V C M, *et al*. An authenticated trust and reputation calculation and management system for cloud and sensor networks integration[J]. IEEE Transactions on Information Forensics and Security, 2015, 10(1): 118-131.
[140] 张国鹏,张海林,赵力强. 基于演化博弈论的移动 Ad Hoc 网络中继协作机制[J]. 控制与决策, 2008, (9): 1077-1080.
[141] 刘凤鸣,丁永生. 基于进化博弈的 P2P 网络中信任计算的动力学分析[J]. 计算机应用研究, 2008, (8): 2460-2462.
[142] 项兴彬,曾国荪,夏冬梅. P2P 环境下文件共享的信任建立博弈模型及稳态分析[J]. 计算机应用研究, 2010, (9): 3496-3499.
[143] Niyato D, Hossain E. Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach[J]. IEEE Transactions on Vehicular Technology, 2009, 58(4): 2008-2017.
[144] Tembine H, Altman E, El-Azouzi R, *et al*. Evolutionary games in wireless networks[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2010, 40(3): 634-646.
[145] Komathy K, Narayanasamy P. Trust-based evolutionary game model assisting AODV routing against selfishness[J]. Journal of Network and Computer Applications, 2008, 31(4): 446-471.
[146] Anastasopoulos M P, Petraki D K, Kannan R, *et al*. TCP throughput adaptation in WiMax networks using replicator dynamics[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2010, 40(3): 647-655.
[147] Wang B, Liu K J R, Clancy T C. Evolutionary cooperative spectrum sensing game: How to collaborate? [J]. IEEE Transactions on Communications, 2010, 58(3): 890-900.
[148] Wang W, Chatterjee M, Kwiat K. Cooperation in wireless networks with unreliable channels[J]. IEEE Transactions on Communications, 2011, 59(10): 2808-2817.
[149] Chen Z, Qiu Y, Liu J, *et al*. Incentive mechanism for selfish nodes in wireless sensor networks based on evolutionary game[J]. Computers & Mathematics with Applications, 2011, 62(9): 3378-3388.
[150] Lin J, Xiong N, Vasilakos A V, *et al*. Evolutionary game-based data aggregation model for wireless sensor networks[J]. IET Communications, 2011, 5(12): 1691-1697.
[151] Zhao B Q, Lui J C S, Chiu D. A mathematical framework for analyzing adaptive incentive protocols in P2P networks[J]. IEEE/ACM Transactions on Networking, 2012, 20(2): 367-380.
[152] Khan M A, Tembine H, Vasilakos A V. Game dynamics and cost of learning in heterogeneous 4G networks[J]. IEEE Journal on Selected Areas in Communications, 2012, 30(1): 198-213.
[153] 孟宪福,王动. 基于重复博弈和惩戒机制的 P2P 协作激励信誉模型[J]. 计算机辅助设计与图形学学报, 2010, (5): 233-236.
[154] 罗俊海,范明钰. 基于博弈的 MANETs 信任模型研究[J]. 计算机研究与发展, 2008, (10): 1704-1710.
[155] 黄宇,曾国荪,袁禄来. 一种基于完全信息扩展博弈的自动信任协商策略[J]. 微电子学与计算机, 2009, (10): 21-24.
[156] 刘继超,曾国荪,袁禄来. 基于开放网络环境下信任建立的博弈模型[J]. 计算机工程, 2009, (2): 167-169.

[157] 陈晶，杜瑞颖，王丽娜，等. 网络环境下一种基于概率密度的信任博弈模型[J]. 电子学报，2010，(2)：427-433.

[158] 孙玉星，赵燕飞，李娅，等. 基于概率博弈的无线自组网信任推荐激励策略的研究[J]. 计算机科学，2011，(4)：65-71.

[159] 桂劲松，吴敏. 基于信任和服务预测的无线接入服务博弈控制方案[J]. 计算机研究与发展，2012，(2)：231-242.

[160] Wang C, Wang R, Chen H, *et al*. Study of automated trust negotiation mechanism based on cache sequence game in P2P environment[J]. Information Technology Journal, 2011, 10(11): 2014-2023.

[161] Jaramillo J J, Srikant R. A game theory based reputation mechanism to incentivize cooperation in wireless ad hoc networks[J]. Ad Hoc Networks, 2010, 8(4): 416-429.

[162] Mejia M, Peña N, Muñoz J L, *et al*. A game theoretic trust model for on-line distributed evolution of cooperation in MANETs[J]. Journal of Network and Computer Applications, 2011, 34(1): 39-51.

[163] Yahyaoui H. A trust-based game theoretical model for Web services collaboration[J]. Knowledge-Based Systems, 2012, 27: 162-169.

[164] 杨东巍，谢福鼎，张永. 基于信任的无线传感器网络时隙分配博弈分析[J]. 计算机工程与设计，2011，(4)：1211-1215，1219.

[165] 李紫川，沈士根，曹奇英. 基于反思机制的 WSNs 节点信任演化模型[J]. 计算机应用研究，2014，(5)：1528-1531.

[166] Komathy K, Narayanasamy P. Secure data forwarding against denial of service attack using trust based evolutionary game[C]. Proc. of IEEE Vehicular Technology Conference, 2008: 31-35.

[167] Feng R, Che S, Wang X, *et al*. An incentive mechanism based on game theory for trust management[J]. Security and Communication Networks, 2014, 7(12): 2318-2325.

[168] Duan J, Gao D, Yang D, *et al*. An energy-aware trust derivation scheme with game theoretic approach in wireless sensor networks for IoT applications[J]. IEEE Internet of Things Journal, 2014, 1(1): 58-69.

[169] Guo W Z, Chen J Y, Chen G L, *et al*. Trust dynamic task allocation algorithm with Nash equilibrium for heterogeneous wireless sensor network[J]. Security and Communication Networks, 2015, 8(10): 1865-1877.

[170] Safa H, Artail H, Tabet D. A cluster-based trust-aware routing protocol for mobile ad hoc networks[J]. Wireless Networks, 2010, 16(4): 969-984.

[171] Sun B, Shrestha D, Yan G, *et al*. Self-propagate mal-packets in wireless sensor networks: Dynamics and defense implications[C]. Proc. of GLOBECOM, 2008: 4961-4965.

[172] Yang Y, Zhu S, Cao G. Improving sensor network immunity under worm attacks: A software diversity approach[C]. Proc. of MobiHoc, 2008: 149-158.

[173] De P, Das S K. Epidemic models, algorithms, and protocols in Wireless Sensor and Ad Hoc Networks[M]. Algorithms and Protocols for Wireless Sensor Networks, Boukerche A, Hoboken: John Wiley & Sons, 2008, 51-75.

[174] Kondakci S, Dincer C. Internet epidemiology: Healthy, susceptible, infected, quarantined, and recovered[J]. Security and Communication Networks, 2011, 4(2): 216-238.

[175] Di Fatta G, Blasa F, Cafiero S, *et al*. Fault tolerant decentralised K-Means clustering for asynchronous large-scale networks[J]. Journal of Parallel and Distributed Computing, 2013, 73(3): 317-329.

[176] 唐辉，郭利新. 机会网络中一种增加控制信息的传染病算法[J]. 广东通信技术，2012，(1)：22-24.

[177] Anagnostopoulos C, Sekkas O, Hadjiefthymiades S. An adaptive epidemic information dissemination model for wireless sensor networks[J]. Pervasive and Mobile Computing, 2012, 8(5): 751-763.
[178] 姜庆五,陈启明. 流行病学方法与模型[M]. 上海:复旦大学出版社, 2007.
[179] Wang Y, Wang J, Zhang L. Cross diffusion-induced pattern in an SI model[J]. Applied Mathematics and Computation, 2010, 217(5): 1965-1970.
[180] 原存德,胡宝安. 具有阶段结构的 SI 传染病模型[J]. 应用数学学报, 2002, (2): 193-203.
[181] Ji C, Jiang D, Shi N. The behavior of an SIR epidemic model with stochastic perturbation[J]. Stochastic Analysis and Applications, 2012, 30(5): 755-773.
[182] 张梅,张凤琴. 一类具有不同感染率的 SIR 模型的稳定性分析[J]. 数学的实践与认识, 2010, (14): 232-236.
[183] 黄娜,黄健民. 具有两种病毒的脉冲时滞传染病 SEIR 模型的研究[J]. 广西师范学院学报(自然科学版), 2011, (4): 28-31.
[184] Yang Q, Jiang D, Shi N, *et al*. The ergodicity and extinction of stochastically perturbed SIR and SEIR epidemic models with saturated incidence[J]. Journal of Mathematical Analysis and Applications, 2012, 388(1): 248-271.
[185] Shen S, Li H, Han R, *et al*. Differential game-based strategies for preventing malware propagation in wireless sensor networks[J]. IEEE Transactions on Information Forensics and Security, 2014, 9(11): 1962-1973.
[186] 王蕊,冯登国,杨轶,等. 基于语义的恶意代码行为特征提取及检测方法[J]. 软件学报, 2012, (2): 378-393.
[187] Egele M, Scholte T, Kirda E, *et al*. A survey on automated dynamic malware-analysis techniques and tools[J]. ACM Computing Surveys, 2012, Vol. 44, Article ID 6, 42 pages.
[188] Santos I, Brezo F, Sanz B, *et al*. Using opcode sequences in single-class learning to detect unknown malware[J]. IET Information Security, 2011, 5(4): 220-227.
[189] 苗甫,王振兴,张连成. 基于流量统计指纹的恶意代码检测模型[J]. 计算机工程, 2011, (18): 131-133.
[190] 孔德光,谭小彬,奚宏生,等. 提升多维特征检测迷惑恶意代码[J]. 软件学报, 2011, (3): 522-533.
[191] 王蕊,苏璞睿,杨轶,等. 一种抗混淆的恶意代码变种识别系统[J]. 电子学报, 2011, (10): 2322-2330.
[192] 张鹏涛,王维,谭营. 基于带有惩罚因子的阴性选择算法的恶意程序检测模型[J]. 中国科学:信息科学, 2011, (7): 798-812.
[193] Dube T, Raines R, Peterson G, *et al*. Malware target recognition via static heuristics[J]. Computers and Security, 2012, 31(1): 137-147.
[194] Chen Z, Roussopoulos M, Liang Z, *et al*. Malware characteristics and threats on the internet ecosystem[J]. Journal of Systems and Software, 2012, 85(7): 1650-1672.
[195] Perdisci R, Ariu D, Giacinto G. Scalable fine-grained behavioral clustering of HTTP-based malware[J]. Computer Networks, 2013, 57(2): 487-500.
[196] Chandramohan M, Tan H B K. Detection of mobile malware in the wild[J]. Computer, 2012, 45(9): 65-71.
[197] 李鹏,王汝传. 基于自相似特性的恶意代码动态分析技术[J]. 南京邮电大学学报(自然科学版), 2012, (3): 86-90.
[198] 李鹏,王汝传,武宁. 基于空间关系特征的未知恶意代码自动检测技术研究[J]. 计算机研究与发展, 2012, (5): 949-957.
[199] 左黎明,汤鹏志,刘二根,等. 基于行为特征的恶意代码检测方法[J]. 计算机工程, 2012, (2):

129-131.
[200] 苗启广，王蕴，曹莹，等. 面向最小行为的恶意程序检测研究[J]. 系统工程与电子技术，2012，(8)：1735-1740.
[201] 陈丹伟，唐平，周书桃. 基于沙盒技术的恶意程序检测模型[J]. 计算机科学，2012，(6A)：12-14.
[202] Peng S，Yu S，Yang A. Smartphone malware and its propagation modeling：A survey[J]. IEEE Communications Surveys and Tutorials，2014，16(2)：925-941.
[203] 王长广，沈玉龙，马建峰. 一种蓝牙环境下恶意程序的传播模型[J]. 西安电子科技大学学报，2009，(1)：94-98.
[204] 李婵婵，蒋国平，宋玉蓉. 动态小世界社团网络上的病毒传播研究[J]. 复杂系统与复杂性科学，2014，11(3)：33-39.
[205] 左焘，宋玉蓉. 考虑连边保护的自适应网络病毒传播模型[J]. 南京邮电大学学报(自然科学版)，2014，34(6)：94-100.
[206] 林昭文，苏飞，马严. 物联网恶意代码传播模型研究(英文)[J]. 中国通信，2011，(1)：79-86.
[207] 徐小龙，熊婧夷，程春玲，等. 一种 P2P 网络恶意代码 4 状态被动传播模型[J]. 解放军理工大学学报(自然科学版)，2011，(6)：582-587.
[208] Ramachandran K K，Sikdar B. Dynamics of malware spread in decentralized peer-to-peer networks [J]. IEEE Transactions on Dependable and Secure Computing，2011，8(4)：617-623.
[209] Shan Z，Wang X，Chiueh T C. Enforcing mandatory access control in commodity os to disable malware[J]. IEEE Transactions on Dependable and Secure Computing，2012，9(4)：541-555.
[210] Peng S，Wang G，Yu S. Modeling the dynamics of worm propagation using two-dimensional cellular automata in smartphones[J]. Journal of Computer and System Sciences，2013，79(5)：586-595.
[211] Song Y R，Jiang G P，Gong Y W. Epidemic propagation on adaptive coevolutionary networks with preferential local-world reconnecting strategy[J]. Chinese Physics B，2013，Vol. 22，Article ID 040205，7 pages.
[212] Yu S，Gu G，Barnawi A，*et al*. Malware propagation in large-scale networks [J]. IEEE Transactions on Knowledge and Data Engineering，2015，27(1)：170-179.
[213] Feng L，Liao X，Han Q，*et al*. Dynamical analysis and control strategies on malware propagation model[J]. Applied Mathematical Modelling，2013，37(16-17)：8225-8236.
[214] Khosroshahy M，Mehmet Ali M K，Qiu D. The SIC botnet lifecycle model：A step beyond traditional epidemiological models[J]. Computer Networks，2013，57(2)：404-421.
[215] Adu-Gyamfi D，Wang Y，Zhang F，*et al*. A model for spreading behavior of passive worms in mobile social networks [J]. Journal of Computational Information Systems，2014，10 (7)：2667-2675.
[216] Wen S，Zhou W，Zhang J，*et al*. Modeling propagation dynamics of social network worms[J]. IEEE Transactions on Parallel and Distributed Systems，2013，24(8)：1633-1643.
[217] Bose A，Shin K G. Agent-based modeling of malware dynamics in heterogeneous environments[J]. Security and Communication Networks，2013，6(12)：1576-1589.
[218] Faghani M R，Nguyen U T. A study of xss worm propagation and detection mechanisms in online social networks[J]. IEEE Transactions on Information Forensics and Security，2013，8(11)：1815-1826.
[219] Wang J，Liu Y，Deng K. Modelling and simulating worm propagation in static and dynamic traffic [J]. IET Intelligent Transport Systems，2014，8(2)：155-163.
[220] Karyotis V，Papavassiliou S. Macroscopic malware propagation dynamics for complex networks with churn[J]. IEEE Communications Letters，2015，19(4)：577-580.

[221] Cheng S M, Ao W C, Chen P Y, *et al*. On modeling malware propagation in generalized social networks[J]. IEEE Communications Letters, 2011, 15(1): 25-27.

[222] Lu Z, Wang W, Wang C. How can botnets cause storms? Understanding the evolution and impact of mobile botnets[C]. Proc. of IEEE INFOCOM, 2014: 1501-1509.

[223] Peng S, Wu M, Wang G, *et al*. Propagation model of smartphone worms based on semi-Markov process and social relationship graph[J]. Computers and Security, 2014, 44: 92-103.

[224] Peng S, Wu M, Wang G, *et al*. Containing smartphone worm propagation with an influence maximization algorithm[J]. Computer Networks, 2014, 74: 103-113.

[225] 付帅,王长广,马建峰. 无线传感器网络中恶意程序的传播模型[J]. 计算机工程, 2011, (3): 129-131.

[226] 王小明,李成博,李英姝. 移动无线传感网恶意数据包传播随机模型[J]. 电子与信息学报, 2013, 35(6): 1290-1297.

[227] Wang X, He Z, Zhao X, *et al*. Reaction-diffusion modeling of malware propagation in mobile wireless sensor networks[J]. Science China Information Sciences, 2013, 56(9): 1-18.

[228] Wang X, He Z, Zhang L. A pulse immunization model for inhibiting malware propagation in mobile wireless sensor networks[J]. Chinese Journal of Electronics, 2014, 23(4): 810-815.

[229] Giannetsos T, Dimitriou T, Krontiris I, *et al*. Arbitrary code injection through self-propagating worms in Von Neumann architecture devices[J]. Computer Journal, 2010, 53(10): 1576-1593.

[230] Khayam S A, Radha H. Using signal processing techniques to model worm propagation over wireless sensor networks[J]. IEEE Signal Processing Magazine, 2006, 23(2): 164-169.

[231] Yanmaz E. Epidemic propagation in overlaid wireless networks[C]. Proc. of GLOBECOM, 2008: 143-147.

[232] De P, Liu Y, Das S K. An epidemic theoretic framework for vulnerability analysis of broadcast protocols in wireless sensor networks[J]. IEEE Transactions on Mobile Computing, 2009, 8(3): 413-425.

[233] De P, Liu Y, Das S K. Deployment-aware modeling of node compromise spread in wireless sensor networks using epidemic theory[J]. ACM Transactions on Sensor Networks, 2009, Vol. 5, Article ID 23, 33 pages.

[234] Mishra B K, Jha N. SEIQRS model for the transmission of malicious objects in computer network [J]. Applied Mathematical Modelling, 2010, 34(3): 710-715.

[235] Wang X, Li Y. An improved SIR model for analyzing the dynamics of worm propagation in wireless sensor networks[J]. Chinese Journal of Electronics, 2009, 18(1): 8-12.

[236] Wang X, Li Q, Li Y. EiSIRS: A formal model to analyze the dynamics of worm propagation in wireless sensor networks[J]. Journal of Combinatorial Optimization, 2010, 20(1): 47-62.

[237] Tang S. A modified SI epidemic model for combating virus spread in Wireless Sensor Networks[J]. International Journal of Wireless Information Networks, 2011, 18(4): 319-326.

[238] Cao H, Ertin E, Arora A. MiniMax equilibrium of networked differential games[J]. ACM Transactions on Autonomous and Adaptive Systems, 2008, Vol. 3, Article ID 14, 21 pages.

[239] Miao X, Zhou X, Wu H. A cooperative differential game model based on network throughput and energy efficiency in wireless networks[J]. Optimization Letters, 2010, 6(1): 55-68.

[240] Lin L, Wang A, Zhou X, *et al*. Noncooperative differential game based efficiency-aware traffic assignment for multipath routing in CRAHN[J]. Wireless Personal Communications, 2012, 62(2): 443-454.

[241] Zhu K, Niyato D, Wang P. Optimal bandwidth allocation with dynamic service selection in heterogeneous wireless networks[C]. Proc. of IEEE Global Telecommunications Conference,

2010: 1-5.

[242] Gu D. A game theory approach to target tracking in sensor networks[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2011, 41(1): 2-13.

[243] Xu H, Zhou X. A Non-cooperative Differential Game Model for Frequency Reuse Based Channel Allocation in Satellite Networks[J]. Wireless Personal Communications, 2014, 79(1): 405-416.

[244] Hu J, Xie Y. A Stochastic Differential Game Theoretic Study of Multipath Routing in Heterogeneous Wireless Networks[J]. Wireless Personal Communications, 2014, 80(3): 971-991.

[245] Theodorakopoulos G, Baras J S, Le Boudec J. Dynamic Network Security Deployment Under Partial Information[C]. Proc. of 46th Annual Allerton Conference on Communication, Control, and Computing, 2008: 261-267.

[246] Omic J, Orda A, Van Mieghem P. Protecting against network infections: A game theoretic perspective[C]. Proc. of IEEE INFOCOM, 2009: 1485-1493.

[247] Bensoussan A, Kantarcioglu M, Hoe S. A game-theoretical approach for finding optimal strategies in a botnet defense model[C]. Proc. of GameSec, 2010: 135-148.

[248] Khouzani M H R, Sarkar S, Altman E. Saddle-point strategies in malware attack[J]. IEEE Journal on Selected Areas in Communications, 2012, 30(1): 31-43.

[249] Aumann R J, Hart S. Handbook of Game Theory with Economic Applications[M]. Amsterdam: Elsevier Press, 1994.

[250] Sterbenz J P G, Hutchison D, çetinkaya E K, *et al*. Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines[J]. Computer Networks, 2010, 54(8): 1245-1265.

[251] Al-Kuwaiti M, Kyriakopoulos N, Hussein S. A comparative analysis of network dependability, fault-tolerance, reliability, security, and survivability[J]. IEEE Communications Surveys and Tutorials, 2009, 11(2): 106-124.

[252] Shen S, Han R, Guo L, *et al*. Survivability evaluation towards attacked WSNs based on stochastic game and continuous-time Markov chain[J]. Applied Soft Computing, 2012, 12(5): 1467-1476.

[253] 王良民,廖闻剑. 无线传感器网络可生存理论与技术研究[M]. 北京:人民邮电出版社, 2011.

[254] Knight J C, Sullivan K J. On the definition of survivability[R]. Department of Computer Science, University of Virginia, Technical Report CS-TR-33-00, 2000.

[255] 杨超,马建峰. 可生存网络系统的形式化定义[J]. 网络安全技术与应用, 2004, (7): 39-41.

[256] Habib M F, Tornatore M, Dikbiyik F, *et al*. Disaster survivability in optical communication networks[J]. Computer Communications, 2013, 36(6): 630-644.

[257] Albano W A, Nogueira M, De Souza J N. A taxonomy for resilience in vehicular Ad hoc networks [J]. IEEE Latin America Transactions, 2015, 13(1): 228-234.

[258] Khan S A, Daachi B, Djouani K. Application of fuzzy inference systems to detection of faults in wireless sensor networks[J]. Neurocomputing, 2012, 94: 111-120.

[259] Khedr A M, Osamy W. Mobility-assisted minimum connected cover in a wireless sensor network [J]. Journal of Parallel and Distributed Computing, 2012, 72(7): 827-837.

[260] Konstantopoulos C, Pantziou G, Gavalas D, *et al*. A rendezvous-based approach enabling energy-efficient sensory data collection with mobile Sinks [J]. IEEE Transactions on Parallel and Distributed Systems, 2012, 23(5): 809-817.

[261] Pu J, Gu Y, Zhang Y, *et al*. A hole-tolerant redundancy scheme for wireless sensor networks[J]. International Journal of Distributed Sensor Networks, 2012, Vol. 2012, Article ID 320108, 10 pages.

[262] 唐林俊. 无线传感网络中部分覆盖与拟连通冗余节点的研究[J]. 传感技术学报, 2011, (6):

895-899.

[263] Banimelhem O, Khasawneh S. GMCAR: Grid-based multipath with congestion avoidance routing protocol in wireless sensor networks[J]. Ad Hoc Networks, 2012, 10(7): 1346-1361.

[264] Xu H, Huang L, Qiao C, *et al*. Bandwidth-power aware cooperative multipath routing for wireless multimedia sensor networks[J]. IEEE Transactions on Wireless Communications, 2012, 11(4): 1532-1543.

[265] Priya S M, Karthikeyan S. An efficient clustered multipath routing to improve lifespan in WSN[J]. International Journal of Computer Science, 2012, 9(2): 182-187.

[266] Rezaie A R, Mirnia M. CMQ: Clustering based multipath routing algorithm to improving QoS in wireless sensor networks[J]. International Journal of Computer Science Issues, 2012, 9(3): 156-160.

[267] Cintron F J, Pongaliur K, Mutka M W, *et al*. Leveraging height in a jumping sensor network to extend network coverage[J]. IEEE Transactions on Wireless Communications, 2012, 11(5): 1840-1849.

[268] Wei L, Qin Z. On-line bi-objective coverage hole healing in hybrid wireless sensor networks[J]. Journal of Computational Information Systems, 2012, 8(13): 5649-5658.

[269] 刘志,裘正定. 基于准格型策略的无线传感网协作覆盖算法[J]. 电子与信息学报, 2010, (10): 2531-2535.

[270] 蒋丽萍,王良民,熊书明,等. 基于感知概率的无线传感器网络 k 重覆盖算法[J]. 计算机应用研究, 2009, (9): 3484-3486.

[271] 沙超,王汝传,黄海平,等. 多媒体传感网中一种基于优化覆盖及压缩代价评估的节能机制[J]. 电子学报, 2011, (10): 2353-2358.

[272] 张万松,王立松. WSN 操作系统可生存性技术的研究与实现[J]. 计算机工程, 2009, (18): 122-124.

[273] Al-Kofahi O M, Kamal A E. Survivability strategies in multihop wireless networks[J]. IEEE Wireless Communications, 2010, 17(5): 71-80.

[274] 李姗姗,廖湘科,朱培栋,等. 基于网络编码的无线传感网多路径传输方法[J]. 软件学报, 2008, (10): 2638-2647.

[275] 赵炜,唐振民,纪淑标,等. 基于网络编码的传感网多径路由模型分析[J]. 计算机工程与设计, 2012, (3): 875-879.

[276] Bari A, Jaekel A, Jiang J, *et al*. Design of fault tolerant wireless sensor networks satisfying survivability and lifetime requirements[J]. Computer Communications, 2012, 35(3): 320-333.

[277] 王良民,郭渊博,詹永照. 容忍入侵的无线传感器网络模糊信任评估模型[J]. 通信学报, 2010, (12): 37-44.

[278] Wang L, Jiang T, Zhu X. Updatable key management scheme with intrusion tolerance for unattended wireless sensor network[C]. Proc. of 54th Annual IEEE Global Telecommunications Conference, 2011: 1-5.

[279] Nabizadeh H, Abbaspour M. IFRP: An intrusion/fault tolerant routing protocol for increasing resiliency and reliability in wireless sensor networks[C]. Proc. of 2011 International Conference on Selected Topics in Mobile and Wireless Networking, 2011: 24-29.

[280] 王海泉,马心意,夏春和. 一种 MANET 可生存性模型的建模方法[J]. 信息安全与通信保密, 2010, (1): 88-92.

[281] 刘宏月,马建峰,王超. 基于容错 CORBA 的可生存网络应用模型[J]. 华中科技大学学报(自然科学版), 2010, (10): 26-30.

[282] 黄继鹏,江建慧. 一种瞬态可生存性评估模型[J]. 内蒙古大学学报(自然科学版), 2011, (5):

481-487.
[283] Ma Z S. Frailty modelling for risk analysis in network security and survivability[J]. International Journal of Information and Computer Security, 2011, 4(3): 276-294.
[284] Ghazisaidi N, Scheutzow M, Maier M. Survivability analysis of next-generation passive optical networks and fiber-wireless access networks[J]. IEEE Transactions on Reliability, 2011, 60(2): 479-492.
[285] Queiroz C, Mahmood A, Tari Z. A probabilistic model to predict the survivability of SCADA systems[J]. IEEE Transactions on Industrial Informatics, 2013, 9(4): 1975-1985.
[286] Zhao C, Liu Y, Yu Z. Assessment of the survivability of networked system based on improved TOPSIS[C]. Proc. of Advanced Research on Computer Science and Information Engineering Communications in Computer and Information Science, 2011: 355-360.
[287] Yi Z, Dohi T. Quantitative comparison of survivability models for wireless ad hoc networks[C]. Proc. of 2nd International Conference on Networking and Computing, 2011: 284-287.
[288] 沈建春,陈佳庆,王志刚. 基于模糊综合评价的信息网络系统可生存性评估[J]. 信息系统工程, 2011, (12): 105-107.
[289] 伍文,孟相如,马志强,等. 基于组合赋权的网络可生存性模糊综合评估[J]. 系统工程与电子技术, 2013, (4): 786-790.
[290] 刘延华,陈国龙,吴瑞芬. 基于云模型和 AHP 的网络信息系统可生存性评估[J]. 通信学报, 2014, (8): 107-115.
[291] Zhao C, Yu Z. Quantitative analysis of survivability based on intrusion scenarios[J]. Lecture Notes in Electrical Engineering, 2012, 140: 701-705.
[292] Di Pietro R, Verde N V. Epidemic theory and data survivability in unattended wireless sensor networks: Models and gaps[J]. Pervasive and Mobile Computing, 2013, 9(4): 588-597.
[293] Bahi J M, Guyeux C, Hakem M, *et al*. Epidemiological approach for data survivability in unattended wireless sensor networks[J]. Journal of Network and Computer Applications, 2014, 46: 374-383.
[294] 张勇实,张乐君,张健沛,等. 基于 QoS 关联分析的分布式系统可生存性评估[J]. 电子科技大学学报, 2013, (01): 109-114.
[295] 王鹏飞,赵文涛,张帆,等. 网络系统可生存能力量化评估的指标体系研究[J]. 计算机工程与科学, 2014, (06): 1050-1056.
[296] 熊琦,王丽娜,刘陶,等. 面向容侵系统可生存性量化的随机博弈模型研究[J]. 小型微型计算机系统, 2008, (10): 1794-1798.
[297] 谢波,肖晓强,徐明,等. 一种基于马尔可夫链的车辆自组网可生存性模型[J]. 计算机应用, 2008, (10): 2577-2579.
[298] Jindal V, Dharmaraja S, Trivedi K S. Markov modeling approach for survivability analysis of cellular networks[J]. International Journal of Performability Engineering, 2011, 7(5): 429-440.
[299] Wang J, Yu Z. Research on quantitative analysis model of MANET survivability[C]. Proc. of 2nd Annual Conference on Electrical and Control Engineering, 2011: 2506-2510.
[300] Parvin S, Hussain F K, Park J S, *et al*. A survivability model in wireless sensor networks[J]. Computers & Mathematics with Applications, 2012, 64(12): 3666-3682.
[301] 赵二虎,阳小龙,彭云峰,等. CPSM:一种增强 IP 网络生存性的客户端主动服务漂移模型[J]. 电子学报, 2010, (9): 2134-2139.
[302] 伍文,孟相如,刘芸江,等. 基于连续时间 Markov 的网络可生存性建模与量化[J]. 吉林大学学报(工学版), 2013, (5): 1395-1400.
[303] Dharmaraja S, Jindal V, Varshney U. Reliability and survivability analysis for UMTS networks:

An analytical approach[J]. IEEE Transactions on Network and Service Management, 2008, 5(3): 132-142.

[304] 林闯,王元卓,杨扬,等. 基于随机 Petri 网的网络可信赖性分析方法研究[J]. 电子学报, 2006, (2): 322-332.

[305] 孙显军,朱亮,高志民,等. 应用随机 Petri 网分析分布式信息系统可生存性[J]. 系统仿真学报, 2008, (S2): 181-186.

[306] 张慧敏,古天龙. 基于 Petri 网模型的 Ad Hoc 网络可生存性分析[J]. 系统仿真学报, 2008, (9): 2487-2490.

[307] 刘密霞,张玉清,洪毅. 基于模糊推理的网络可生存性的建模与分析[J]. 通信学报, 2009, (1): 31-37.

[308] 赵明峰,周亚建,任东晓,等. 基于 SPN 的 IMS 系统可生存性建模及分析[J]. 应用科学学报, 2012, (3): 239-244.

[309] 刘梅霞,古天龙. 基于 GSPN 模型的 Ad Hoc 网络可生存性分析[J]. 桂林电子科技大学学报, 2009, (2): 82-87.

[310] 李良斌,王劲林,陈君. 基于着色 Petri 网的系统可生存性仿真平台[J]. 计算机工程, 2012, (2): 14-16.

[311] 王元卓,林闯,程学旗,等. 基于随机博弈模型的网络攻防量化分析方法[J]. 计算机学报, 2010, (9): 1748-1762.

[312] Buzacott J A. Markov approach to finding failure times of repairable systems [J]. IEEE Transactions on Reliability, 1970, R-19(4): 128-134.

[313] Sallhammar K, Helvik B E, Knapskog S J. On stochastic modeling for integrated security and dependability evaluation[J]. Journal of Networks, 2006, 1(5): 31-42.

[314] Xing F, Wang W. On the survivability of wireless Ad Hoc networks with node misbehaviors and failures[J]. IEEE Transactions on Dependable and Secure Computing, 2010, 7(3): 284-299.

[315] Peng S, Jia W, Wang G. Survivability evaluation in large-scale mobile Ad-Hoc networks[J]. Journal of Computer Science & Technology, 2009, 24(4): 761-774.

[316] Chen D, Garg S, Trivedi K S. Network survivability performance evaluation: a quantitative approach with applications in wireless ad-hoc networks[C]. Proc. of the 5th ACM International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems, 2002: 61-68.

[317] Sedaghatbaf A, Abdollahi Azgomi M. Attack modelling and security evaluation based on stochastic activity networks[J]. Security and Communication Networks, 2014, 7(4): 714-737.

[318] 肖宇峰,陈山枝,李昕,等. 用 OBDD 算法评估无线传感网的可靠度和结点重要性[J]. 高技术通讯, 2009, (12): 1245-1250.

[319] Shrestha A, Xing L, Liu H. Modeling and evaluating the reliability of wireless sensor networks [C]. Proc. of Annual Reliability and Maintainability Symposium, 2007: 186-191.

[320] Xiao Y F, Chen S Z, Li X, *et al*. Reliability evaluation of wireless sensor networks using an enhanced OBDD algorithm[J]. Journal of China Universities of Posts and Telecommunications, 2009, 16(5): 62-70.

[321] Bruneo D, Distefano S, Longo F, *et al*. Reliability assessment of wireless sensor nodes with non-linear battery discharge[C]. Proc. of 2010 IFIP Wireless Days, 2010: 1-5.

[322] Wu H, Wang Y, Dang H, *et al*. Analytic, simulation, and empirical evaluation of delay/fault-tolerant mobile sensor networks[J]. IEEE Transactions on Wireless Communications, 2007, 6(9): 3287-3296.

[323] 肖坤,古天龙,常亮. 一种 Ad Hoc 网络可生存性度量方法[J]. 桂林电子科技大学学报, 2011, (3): 213-216.

[324] 肖志力，何明，肖登海，等. 网络信息系统的可生存性评估研究[J]. 计算机工程与应用，2009，(14)：18-21.

[325] 魏昭，夏春和，何冰，等. 一种移动 Ad Hoc 网络可生存性模型建模及仿真验证方法[J]. 计算机学报，2013，(7)：1465-1474.

[326] Kim K, Roh B H, Ko Y B, *et al*. Survivability measure for multichannel MANET-based tactical networks[C]. Proc. of 13th International Conference on Advanced Communication Technology: Smart Service Innovation through Mobile Interactivity, 2011: 1049-1053.

[327] Wang C, Ming L, Zhao J, *et al*. A general framework for network survivability testing and evaluation[J]. Journal of Networks, 2011, 6(6): 831-841.

[328] Wang J L, Yu Z W. Research on evaluation of the MANET system survivability[J]. Procedia Environmental Sciences, 2011, 10(PART A): 51-57.

[329] Ming L, Huang M, Wang D, *et al*. Research on survivability metrics based on survivable process of network system[C]. Proc. of 4th International Conference on Security of Information and Networks, 2011: 247-250.

[330] Sterbenz J P G, çetinkaya E K, Hameed M A, *et al*. Evaluation of network resilience, survivability, and disruption tolerance: Analysis, topology generation, simulation, and experimentation[J]. Telecommunication Systems, 2013, 52(2): 705-736.

[331] Lin F Y S, Chen P Y, Chen Q T. Near-optimal evaluation of network survivability under multi-stage attacks[J]. Lecture Notes in Computer Science, 2012, 7345: 391-399.

[332] 吴庆涛，刘彬，郑瑞娟，等. 自律入侵容忍系统的可生存性评估[J]. 计算机工程，2012，(5)：114-116.

[333] Chang C, Zhu C, Wang H, *et al*. Survivability evaluation of cluster-based wireless sensor network under DoS attack[J]. Communications in Computer and Information Science, 2012, 312: 126-132.

[334] 马驰，李陟，张宏，等. 针对节点失效的 MANET 路由抗毁研究[J]. 计算机研究与发展，2012，(3)：550-557.

[335] Rak J. Measures of region failure survivability for wireless mesh networks[J]. Wireless Networks, 2014, 21(2): 673-684.

[336] Lye K, Wing J M. Game strategies in network security[J]. International Journal of Information Security, 2005, 4(1): 71-86.

[337] Chen G, Shen D, Kwan C, *et al*. Game theoretic approach to threat prediction and situation awareness[J]. Journal of Advances in Information Fusion, 2007, 2(1): 1-14.

[338] Liu D, Wang X, Camp J. Game-theoretic modeling and analysis of insider threats[J]. International Journal of Critical Infrastructure Protection, 2008, 1: 75-80.

[339] Nguyen K C, Alpcan T, Başar T. Stochastic games for security in networks with interdependent nodes[C]. Proc. of the 2009 International Conference on Game Theory for Networks, 2009: 697-703.

[340] Fu F, van der Schaar M. Learning to compete for resources in wireless stochastic games[J]. IEEE Transactions on Vehicular Technology, 2009, 58(4): 1904-1919.

[341] Niyato D, Wang P, Hossain E, *et al*. Exploiting mobility diversity in sharing wireless access: A game theoretic approach[J]. IEEE Transactions on Wireless Communications, 2010, 9(12): 3866-3877.

[342] Wang B, Wu Y, Liu K J R, *et al*. An anti-jamming stochastic game for cognitive radio networks [J]. IEEE Journal on Selected Areas in Communications, 2011, 29(4): 877-889.

[343] 沈昌祥，张焕国，冯登国，等. 信息安全综述[J]. 中国科学(E 辑：信息科学)，2007，(2)：129-150.

[344] Parvin S, Kim D S, Lee S M, *et al*. Achieving availability and survivability in wireless sensor

networks by software rejuvenation[C]. Proc. of the 4th International Workshop on Security, Privacy and Trust in Pervasive and Ubiquitous Computing, 2008: 13-18.

[345] Korkmaz T, Sarac K. Characterizing link and path reliability in large-scale wireless sensor networks [C]. Proc. of 2010 IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications, 2010: 217-224.

[346] 何明,董强,袁黎苗,等. 无线传感器网络的可靠性评估模型[J]. 解放军理工大学学报(自然科学版), 2010, (4): 392-396.

[347] 王良民,马建峰,王超. 无线传感器网络拓扑的容错度与容侵度[J]. 电子学报, 2006, (8): 1446-1451.

[348] 詹永照,饶静宜,王良民. 基于攻击效果的 WSN 路由安全评估模型[J]. 计算机科学, 2010, (7): 70-73.

[349] Masoum A, Jahangir A, Taghikhaki Z, *et al*. Survivability modeling of wireless sensor networks [C]. Proc. of the 2008 IEEE International Symposium on Wireless Communication Systems, 2008: 593-597.

[350] 王海涛,朱世才,陈晖,等. 应急通信中基于 ANP 的 WSN 可生存性评价指标体系研究[J]. 传感技术学报, 2014, (4): 557-563.

[351] 朱世才,王海涛,吴连才,等. 基于 SMP 的分簇 WSN 生存性评估模型[J]. 传感技术学报, 2014, (3): 383-387.

[352] Ma Z S, Krings A W. Insect population inspired wireless sensor networks: A unified architecture with survival analysis, evolutionary game theory, and hybrid fault models [C]. Proc. of International Conference on BioMedical Engineering and Informatics, 2008: 636-643.

[353] Ma Z S, Krings A W. Dynamic hybrid fault modeling and extended evolutionary game theory for reliability, survivability and fault tolerance analyses[J]. IEEE Transactions on Reliability, 2011, 60(1): 180-196.

[354] Petridou S, Basagiannis S, Roumeliotis M. Survivability analysis using probabilistic model checking: A study on wireless sensor networks[J]. IEEE Systems Journal, 2013, 7(1): 4-12.

[355] Yang H, Qin Y, Feng G, *et al*. Online monitoring of geological CO_2 storage and leakage based on wireless sensor networks[J]. IEEE Sensors Journal, 2013, 2(13): 556-562.

[356] Torfs T, Sterken T, Brebels S, *et al*. Low power wireless sensor network for building monitoring [J]. IEEE Sensors Journal, 2013, 13(3): 909-915.

[357] Suryadevara N K, Mukhopadhyay S C. Wireless sensor network based home monitoring system for wellness determination of elderly[J]. IEEE Sensors Journal, 2012, 12(6): 1965-1972.

[358] Tiliute D E. Security of mobile ad hoc wireless networks: a brief survey[J]. Advances in Electrical and Computer Engineering, 2007, 7(2): 37-40.

[359] Mpitziopoulos A, Gavalas D, Konstantopoulos C, *et al*. A survey on jamming attacks and countermeasures in WSNs[J]. IEEE Communications Surveys and Tutorials, 2009, 11(4): 42-56.

[360] Tang C, Wu D. An efficient mobile authentication scheme for wireless networks[J]. IEEE Transactions on Wireless Communications, 2008, 7(4): 1408-1416.

[361] Liu J, Yue G, Shen S, *et al*. A Game-Theoretic Response Strategy for Coordinator Attack in Wireless Sensor Networks[J]. Scientific World Journal, 2014, Vol. 2014, Article ID 950618, 10 pages.

[362] Chong C, Kumar S P. Sensor networks: evolution, opportunities, and challenges[J]. Proceedings of the IEEE, 2003, 91(8): 1247-1256.

[363] Perkins C E, Bhagwat P. Highly dynamic Destination-Sequenced Distance-Vector routing (DSDV) for mobile computers[J]. Computer Communication Review, 1994, 24(4): 234-244.

[364] Johnson D B, Maltz D A. Dynamic source routing in ad hoc wireless networks[J]. The Kluwer International Series in Engineering and Computer Science, 1996, 353: 153-181.

[365] Perkins C E, Royer E M. Ad-hoc on-demand distance vector routing[C]. Proc. of the Second IEEE Workshop on Mobile Computer Systems and Applications, 1999: 90-100.

[366] Heinzelman W, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless sensor networks[C]. Proc. of the Hawaii International Conference System Sciences, 2000: 1-10.

[367] Lindsey S, Raghavendra C S. PEGASIS: power efficient gathering in sensor information systems [C]. Proc. of the IEEE Aerospace Conference, 2002: 1125-1130.

[368] Chang J H, Tassiulas L. Maximum lifetime routing in wireless sensor networks[J]. IEEE/ACM Transactions on Networking, 2004, 12(4): 609-619.

[369] Zimmerling M, Dargie W, Reason J M. Energy-efficient routing in linear wireless sensor networks [C]. Proc. of IEEE International Conference on Mobile Ad Hoc and Sensor Systems, 2007: 1-3.

[370] Shao F, Shen X, Cai L. Energy efficient reliable routing in wireless sensor networks[C]. Proc. of First International Conference Communications and Networking, 2006: 1-5.

[371] Chang H, Tassiulas L. Energy conserving routing in wireless ad hoc networks[C]. Proc. of IEEE INFOCOM, 2000: 22-31.

[372] Cuomo F, Abbagnale A, Cipollone E. Cross-layer network formation for energy-efficient IEEE 802.15.4/ZigBee Wireless Sensor Networks[J]. Ad Hoc Networks, 2013, 11(2): 672-686.

[373] Perrig A, Canetti R, Tygar J, *et al*. Efficient authentication and signing of multicast streams over lossy channels[C]. Proc. of IEEE Symposium on Security and Privacy, 2000: 56-73.

[374] Law Y W, Hoesel L V, Doumen J, *et al*. Energy-efficient link-layer jamming attacks against wireless sensor network MAC protocols[C]. Proc. of Third ACM Workshop Security of Ad Hoc and Sensor Networks, 2005: 76-88.

[375] Xu W, Ma K, Trappe W, *et al*. Jamming sensor networks: attack and defense strategies[J]. IEEE Networks, 2006, 20(3): 41-47.

[376] Yao Z, Kim D, Doh Y. PLUS: parameterized and localized trust management scheme for sensor networks security[C]. Proc. of IEEE International Conference on Mobile Ad hoc and Sensor Systems, 2006: 437-446.

[377] Gabrielli A, Mancini L V, Setia S, *et al*. Securing topology maintenance protocols for sensor networks[J]. IEEE Transactions on Dependable and Secure Computing, 2011, 8(3): 450-465.

[378] Vidgren N, Haataja K, Patiño-Andres J L, *et al*. Security threats in ZigBee-enabled systems: Vulnerability evaluation, practical experiments, countermeasures, and lessons learned[C]. Proc. of the Annual Hawaii International Conference on System Sciences, 2013: 5132-5138.

[379] Patel H J, Temple M A, Baldwin R O. Improving ZigBee device network authentication using ensemble decision tree classifiers with radio frequency distinct native attribute fingerprinting[J]. IEEE Transactions on Reliability, 2014, 64(1): 221-233.

[380] Bakhache B, Ghazal J M, El Assad S. Improvement of the security of ZigBee by a new chaotic algorithm[J]. IEEE Systems Journal, 2013, 8(4): 1021-1030.

[381] Choi K, Yun M, Chae K, *et al*. An enhanced key management using ZigBee Pro for wireless sensor networks[C]. Proc. of International Conference on Information Networking, 2012: 399-403.

[382] Rosli R, Yusoff Y M, Hashim H. Performance analysis of ID-based authentication on ZigBee transceiver[C]. Proc. of IEEE Symposium on Wireless Technology and Applications, 2012: 187-191.

[383] Xu Y, Jiang Y, Hu C, *et al*. A balanced security protocol of Wireless Sensor Network for Smart

Home[C]. Proc. of International Conference on Signal Processing Proceedings, 2015: 2324-2327.

[384] Ramsey B W, Temple M A, Mullins B E. PHY foundation for multi-factor ZigBee node authentication[C]. Proc. of GLOBECOM, 2012: 795-800.

[385] Jokar P, Arianpoo N, Leung V C M. Spoofing prevention using received signal strength for ZigBee-based home area networks [C]. Proc. of IEEE International Conference on Smart Grid Communications, 2013: 438-443.

[386] Tseng L C, Chien F T, Zhang D, *et al*. Network selection in cognitive heterogeneous networks using stochastic learning[J]. IEEE Communications Letters, 2013, 17(12): 2304-2307.

[387] Jiang C X, Chen Y, Liu K J R. Distributed adaptive networks: A graphical evolutionary game-theoretic view[J]. IEEE Transaction on Signal Processing, 2013, 61(22): 5675-5688.

[388] Skaperdas S. Contest success functions[J]. Economic Theory, 1996, 7(2): 283-290.

[389] Smith J M. Evolution and the theory of games [M]. Cambridge: Cambridege University Press, 1982.

[390] Rong C, Nguyen S T, Jaatun M G. Beyond lightning: a survey on security challenges in cloud computing[J]. Computers & Electrical Engineering, 2013, 39(1): 47-54.

[391] 徐剑,周福才,陈旭,等. 云计算中基于认证数据结构的数据外包认证模型[J]. 通信学报, 2011, 32(7): 153-160.

[392] Zissis D, Lekkas D. Addressing cloud computing security issues[J]. Future Generation Computer Systems, 2011, 28: 583-592.

[393] Zhao R, Yue C. Toward a secure and usable cloud-based password manager for web browsers[J]. Computers & Security, 2014, 46(3): 32-47.

[394] Khalila I, Khreishahb A, Azeemc M. Consolidated identity management system for secure mobile cloud computing[J]. Computer Networks, 2014, 65(2): 99-110.

[395] Younis Y A, Kifayat K, Merabti M. An access control model for cloud computing[J]. Journal of Information Security and Applications, 2014, 19(1): 45-60.

[396] Ruj S, Stojmenovic M, Nayak A. Decentralized access control with anonymous authentication of data stored in clouds[J]. IEEE Transactions on Parallel and Distributed Systems, 2014, 25(2): 384-394.

[397] Wang H, Wu S, Chen M, *et al*. Security protection between users and the mobile media cloud[J]. IEEE Communications Magazine, 2014, 52(3): 73-79.

[398] 曹洁,曾国荪,姜火文,等. 云环境下服务信任感知的可信动态级调度方法[J]. 通信学报, 2014, 35(11): 39-41.

[399] Yang L, Wang W, Chen Y, *et al*. A privacy-aware framework for online advertisement targeting [C]. Proc. of IEEE Global Communications Conference, 2013: 3145-3150.

[400] Wang C, Wang Q, Ren K, *et al*. Privacy-preserving public auditing for data storage security in cloud computing[C]. Proc. of IEEE INFOCOM, 2010: 1-9.

[401] Wang Q, Wang C, Ren K, *et al*. Enabling public auditability and data dynamics for storage security in cloud computing[J]. IEEE Transactions on Parallel and Distributed Systems, 2011, 22(5): 847-859.

[402] 盛刚,温涛,郭权,等. 云计算中安全的向量点积计算[J]. 东北大学学报, 2013, 34(6): 786-791.

[403] Ma X, Zhang J, Tao J, *et al*. DNSRadar: Outsourcing malicious domain detection based on distributed cache-footprints[J]. IEEE Transactions on Information Forensics and Security, 2014, 9(11): 1906-1921.

[404] Van den Berg E, Zhang T, Pietrowicz S. Blend-in: A privacy-enhancing certificate-selection method for vehicular communication[J]. IEEE Transactions on Vehicular Technology, 2009, 58

(9)：5190-5199.

[405] Wasef A，Jiang Y，Shen X. An efficient distributed-certificate-service scheme for vehicular networks[J]. IEEE Transactions on Vehicular Technology，2010，59(2)：533-549.

[406] 郭晶晶，马建峰，李琦，等. 基于博弈论的移动自组织网络的信任管理方法[J]. 通信学报，2014，35(11)：49-50.

[407] Liu Z，Joy A W，Thompson R A. A dynamic trust model for mobile ad hoc networks[C]. Proc. of 10th IEEE Int. Workshop on Future Trends of Distributed Computing Systems，2004：80-85.

[408] Pirzada A A，McDonald C. Trust establishment in pure ad hoc Networks[J]. Wireless Personal Communications，2006，37(1)：39-168.

[409] Li X，Gui X. Trust quantitative model with multiple decision factors in trusted network[J]. Chinese Journal of Computers，2009，32(3)：405-416.

[410] Li X，Jia Z，Zhang P，*et al*. Trust-based on-demand Multi-path routing in mobile ad hoc networks [J]. IET Information Security，2010，4(4)：212-223.

[411] Xia H，Jia Z，Ju L，*et al*. Trust management model for mobile ad hoc network based on analytic hierarchy process and fuzzy theory[J]. IET Wireless Sensor Systems，2011，1(4)：248-266.

[412] Mui L. Computational models of trust and reputation：agents，evolutionary games，and social networks[D]. MIT，Massachusetts，2003.

[413] Luo J，Liu X，Fan M. A trust model based on fuzzy recommendation for mobile ad hoc networks [J]. Computer Networks，2009，53(14)：2396-2407.

[414] 罗俊海，范明钰. 基于置信度的 MANETs 主观信任管理模型[J]. 计算机研究与发展，2010，47(3)：515-523.

[415] Lia X，Xiong Y，Ma J，*et al*. An efficient and security dynamic identity based authentication protocol for multi-server architecture using smart cards[J]. Journal of Network and Computer Applications，2012，35(2)：763-769.

[416] Khider H，Osman T，Sherkat N. Attribute-based authorization for grid computing[C]. International Conference on Intelligent Systems，Modelling and Simulation，2010：71-74.

[417] Liu C. Cloud service access control system based on ontologies[J]. Advances in Engineering Software，2014，69(3)：26-36.

[418] Liu J，Shen S，Yue G，*et al*. A stochastic evolutionary coalition game model of secure and dependable virtual service in Sensor-Cloud[J]. Applied Soft Computing，2015，30：123-135.

[419] Ristenpart T，Tromer E，Shacham H，*et al*. Hey，you，get off of my cloud：Exploring information leakage in third-party compute clouds[C]. Proc. of the ACM Conference on Computer and Communications Security，2009：199-212.

[420] Zhang Y，Juels A，Oprea A，*et al*. HomeAlone：Co-residency detection in the cloud via side-channel analysis[C]. Proc. of IEEE Symposium on Security and Privacy，2011：313-328.

[421] Santos N，Gummadi K P，Rodrigues R. Towards trusted cloud computing[EB/OL]. http://www.mpi-sws.org/~gummadi/papers/trusted_cloud.pdf.

[422] Liu H. A new form of DoS attack in a cloud and its avoidance mechanism[C]. Proc. of the 2010 ACM workshop on Cloud computing security workshop，2010：65-76.

[423] Girma A，Garuba M，Li J，*et al*. Analysis of DDoS attacks and an introduction of a hybrid statistical model to detect DDoS attacks on cloud computing environment[C]. Proc. of 12th International Conference on Information Technology：New Generations，2015：212-217.

[424] Chouhan V，Peddoju S. Hierarchical storage technique for maintaining hop-count to prevent DoS attack in cloud computing[C]. Proc. of International Conference on Advances in Intelligent Systems and Computing，2012：511-518.

[425] Cong W, Qian W, Kui R, *et al*. Toward secure and dependable storage services in cloud computing [J]. IEEE Transactions on Services Computing, 2012, 5(2): 220-232.

[426] Kim J, Jeong H, Cho I, *et al*. A secure smart-work service model based OpenStack for Cloud computing[J]. Cluster Computing, 2013, 17(3): 691-702.

[427] Arshad J, Azad M A, Jokhio I A, *et al*. Intrusion damage assessment for multi-stage attacks for clouds[J]. IET Communications, 2013, 7(12): 1304-1315.

[428] Zhou C V, Leckie C, Karunasekera S. A survey of coordinated attacks and collaborative intrusion detection[J]. Computers & Security, 2010, 29(1): 124-140.

[429] Kwon H, Kim T, Yu S, *et al*. Self-similarity based lightweight intrusion detection method for cloud computing[C]. Proc. of the Third international conference on Intelligent information and database systems, 2011: 353-362.

[430] Noor T, Sheng Q. Credibility-based trust management for services in cloud environments[J]. Lecture Notes in Computer Science, 2011, 7084: 328-343.

[431] Mui V N, Eui-Nam H. An efficient key management for secure multicast in sensor-cloud[C]. Proc. of First ACIS/JNU International Conference on Computers, Networks, Systems and Industrial Engineering, 2011: 3-9.

[432] Xie Y, Yu S. A large-scale hidden semi-Markov model for anomaly detection on user browsing behaviors[J]. IEEE/ACM Transactions on Networking, 2009, 17(1): 54-65.

[433] Idziorek J, Tannian M. Exploiting cloud utility models for profit and ruin[C]. Proc. of IEEE International Conference on Cloud Computing, 2011: 33-40.

[434] Kholidy H A, Erradi A, Abdelwahed S, *et al*. A finite state hidden Markov model for predicting multistage attacks in cloud systems[C]. Proc. of IEEE 12th International Conference on Dependable, Autonomic and Secure Computing, 2014: 14-19.

[435] Zhang H, Ye L, Du X, *et al*. Protecting private cloud located within public cloud[C]. Proc. of IEEE Global Communications Conference, 2013: 677-681.

[436] Chen C, Guan D J, Huang Y, *et al*. State-based attack detection for cloud[C]. Proc. of 2013 IEEE International Symposium on Next-Generation Electronics, 2013: 177-180.

[437] Fan G, Yu H, Chen L, *et al*. A game theoretic method to model and evaluate attack-defense strategy in cloud computing[C]. Proc. of 2013 IEEE International Conference on Services Computing, 2013: 659-666.

[438] Bedi H, Shiva S. Securing cloud infrastructure against co-resident DoS attacks using game theoretic defense mechanisms[C]. Proc. of International Conference on Advances in Computing, Communications and Informatics, 2012: 463-469.

[439] Varadarajan V, Kooburat T, Farley B, *et al*. Resource-freeing attacks: Improve your cloud performance (at your neighbor's expense)[C]. Proc. of ACM Conference on Computer and Communications Security, 2012: 281-292.

[440] Zhou F F, Goel M, Desnoyers P, *et al*. Scheduler vulnerabilities and coordinated attacks in cloud computing[J]. Journal of Computer Security, 2013, 21(4): 1-35.

[441] Zhang Y, Li M, Bai K, *et al*. Incentive compatible moving target defense against VM-colocation attacks in clouds[J]. Information Security and Privacy Research, 2012, 376: 388-399.

[442] Poor H V. Information and inference in the wireless physical layer[J]. IEEE Wireless Communications, 2012, 19(1): 40-47.

[443] 龙航，袁广翔，王静，等. 物理层安全技术研究现状与展望[J]. 电信科学，2011，27(9)：60-65.

[444] Wyner A D. The wire-tap channel[J]. Bell System Technical Journal, 1975, 54(8): 1355-1387.

[445] Leung-Yan-Cheong S K, Hellman M E. The Gaussian wire-tap channel[J]. IEEE Transactions on

Information Theory, 1978, IT-24(4): 451-456.
[446] Csiszár I, Körner J. Broadcast channels with confidential messages[J]. IEEE Transactions on Information Theory, 1978, IT-24(3): 339-348.
[447] Awan Z H, Zaidi A, Vandendorpe L. Secure communication over parallel relay channel[J]. IEEE Transactions on Information Forensics and Security, 2012, 7(2): 359-371.
[448] Liu R, Marč I, Spasojevč P, *et al*. Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions[J]. IEEE Transactions on Information Theory, 2008, 54(6): 2493-2507.
[449] 吉江,金梁,黄开枝.基于人工噪声的MISO保密容量分析[J].通信学报, 2012, 33(10): 138-142.
[450] Gerbracht S, Scheunert C, Jorswieck E A. Secrecy outage in MISO systems with partial channel information[J]. IEEE Transactions on Information Forensics and Security, 2012, 7(2): 704-716.
[451] Shafiee S, Liu N, Ulukus S. Towards the secrecy capacity of the gaussian MIMO wire-tap channel: The 2-2-1 channel[J]. IEEE Transactions on Information Theory, 2009, 55(9): 4033-4039.
[452] Zhu J, Schober R, Bhargava V K. Secure transmission in multicell massive MIMO systems[J]. IEEE Transactions on Wireless Communications, 2014, 13(9): 4766-4781.
[453] Jiang G, Shen S, Hu K, *et al*. Evolutionary game-based secrecy rate adaptation in wireless sensor networks[J]. International Journal of Distributed Sensor Networks, 2015, Vol. 2015, Article ID 975454, 13 pages.
[454] 沈士根,黄龙军,屠昂燕,等.基于演化博弈的传感节点保密率自适应调节方法[J].电信科学, 2014, 30(11): 73-79.
[455] Mukherjee A, Fakoorian S A A, Huang J, *et al*. Principles of physical layer security in multiuser wireless networks: A survey[J]. IEEE Communications Surveys and Tutorials, 2014, 16(3): 1550-1573.
[456] 李翔宇,金梁,黄开枝,等.基于联合信道特征的中继物理层安全传输机制[J].计算机学报, 2012, (7): 1399-1406.
[457] 王亚东,黄开枝,吉江.一种多天线信道特征投影物理层安全编码算法[J].电子与信息学报, 2012, (7): 1653-1658.
[458] 卫红权,罗文宇,兰巨龙,等.基于扰动理论无线物理层安全模型及敏度分析[J].通信学报, 2013, (6): 201-206.
[459] 陈涛,余华,韦岗.认知无线电网络的物理层安全研究及其鲁棒性设计[J].电子与信息学报, 2012, (4): 770-775.
[460] 罗苗,王慧明,殷勤业.基于协作波束形成的中继阻塞混合无线物理层安全传输[J].中国科学:信息科学, 2013, (4): 445-458.
[461] 李桥龙,金梁.基于最小信息泄露的线性随机化实现物理层安全传输[J].通信学报, 2013, (7): 42-48.
[462] 邓浩,王慧明,王文杰.基于多节点分组协作干扰的无线物理层安全传输[J].中国科学:信息科学, 2014, (11): 1482-1494.
[463] 李明亮,黄开枝,钟州.基于空频联合加扰的物理层安全算法[J].电子与信息学报, 2013, (12): 2966-2971.
[464] 林通,黄开枝,罗文宇.一种基于多载波的多播系统物理层安全方案[J].电子与信息学报, 2013, (6): 1338-1343.
[465] 吉江,刘璐,金梁,等.随机发送参考的多天线系统物理层安全传输算法[J].中国科学:信息科学, 2014, (2): 254-262.
[466] 崔波,刘璐,李翔宇,等.有限字符输入的空间调制物理层安全传输方法[J].通信学报, 2015, (2): 162-171.

[467] 赵耀环,谢梦非,尚勇.物理层安全中的最优中继选择及协同干扰策略[J].电子学报,2015,(4):791-794.

[468] Wang L, Yang N, Elkashlan M, *et al*. Physical layer security of maximal ratio combining in two-wave with diffuse power fading channels[J]. IEEE Transactions on Information Forensics and Security, 2014, 9(2): 247-258.

[469] Hong L, Chen W. Information theory and cryptography based secured communication scheme for cooperative MIMO communication in wireless sensor networks[J]. Ad Hoc Networks, 2014, 14: 95-105.

[470] Hanif M F, Tran L N, Juntti M, *et al*. On linear precoding strategies for secrecy rate maximization in multiuser multiantenna wireless networks[J]. IEEE Transactions on Signal Processing, 2014, 62(14): 3536-3551.

[471] Chae S H, Choi W, Lee J H, *et al*. Enhanced secrecy in stochastic wireless networks: Artificial noise with secrecy protected zone[J]. IEEE Transactions on Information Forensics and Security, 2014, 9(10): 1617-1628.

[472] 肖宛阳,黄开枝,罗兴国,等.基于博弈论的物理层安全建模及现状分析[J].信息工程大学学报,2013,14(4):402-409.

[473] 洪颖,黄开枝,罗文宇,等.一种基于两次报价博弈机制的安全中继选择方法[J].信息工程大学学报,2014,(5):551-556.

[474] 都晨辉,宋梅,王莉,等.基于Stackelberg博弈的协作干扰策略[J].北京邮电大学学报,2014,(5):11-15.

[475] 黄开枝,洪颖,罗文宇,等.基于演化博弈机制的物理层安全协作方法[J].电子与信息学报,2015,(1):193-199.

[476] 林胜斌,黄开枝,王文,等.存在恶意干扰者时一种基于连续零和博弈的物理层安全传输方法[J].信号处理,2015,(6):720-726.

[477] 吕健体,沈士根,马绚,等.基于博弈论的无线传感器网络保密容量优化研究[J].计算机应用与软件,2015,(6):263-266.

[478] Wang A, Cai Y, Yang W, *et al*. Auction-based security game for multiuser cooperative networks[J]. Frequenz, 2013, 67(5-6): 195-202.

[479] Yuksel M, Liu X, Erkip E. A secure communication game with a relay helping the eavesdropper[J]. IEEE Transactions on Information Forensics and Security, 2011, 6(3): 818-830.

[480] Mukherjee A, Swindlehurst A L. Jamming games in the MIMO wiretap channel with an active eavesdropper[J]. IEEE Transactions on Signal Processing, 2013, 61(1): 82-91.

[481] Han Z, Marina N, Debbah M, *et al*. Physical layer security game: Interaction between source, eavesdropper, and friendly jammer[J]. Eurasip Journal on Wireless Communications and Networking, 2010, Vol. 2009, Article ID 452907, 10 pages.

[482] Gabry F, Li N, Schrammar N, *et al*. On the optimization of the secondary transmitter's strategy in cognitive radio channels with secrecy[J]. IEEE Journal on Selected Areas in Communications, 2014, 32(3): 451-463.

[483] Chu Z, Cumanan K, Ding Z, *et al*. Secrecy rate optimizations for a MIMO secrecy channel with a cooperative jammer[J]. IEEE Transactions on Vehicular Technology, 2015, 64(5): 1833-1847.

[484] Qu J, Cai Y, Wu D, *et al*. Stackelberg game based power allocation for physical layer security of device-to-device communication underlaying cellular networks[J]. Frequenz, 2014, 68(5-6): 285-295.

[485] Saad W, Han Z, Başar T, *et al*. Distributed coalition formation games for secure wireless transmission[J]. Mobile Networks and Applications, 2011, 16(2): 231-245.

[486] Fakoorian S A A, Swindlehurst A L. MIMO interference channel with confidential messages: Achievable secrecy rates and precoder design[J]. IEEE Transactions on Information Forensics and Security, 2011, 6(3): 640-649.

[487] Liu Z, Shang Y, Zhang R, *et al*. Relay selection based on coalitional game for secure wireless networks [J]. IET Communications, 2014, 8(8): 1355-1363.

[488] Hou L, Fu X. Physical layer security with dynamic behaviour cooperator based on coalitional game [J]. IET Communications, 2014, 8(8): 1258-1264.

[489] Cressman R. Evolutionary Dynamics and Extensive Form Games [M]. Cambridge: MIT Press, 2003.